ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

ON THE LUSTERNIK-SCHNIRELMANN CATEGORY

BY RALPH H. FOX

(Received August 10, 1939)

Introduction. In their study of the calculus of variations in the large [36; 37],[1] in particular, of the existence of geodesics on a surface of the topological type of the sphere, Lusternik and Schnirelmann were led [38; 39; 40; 42] to introduce new topological invariants—the category and the combinatorial or homology (mod 2) category. The rôle of the homology category was to furnish a medium for the application of homology theory to the calculation of the category of manifolds.

Although the category occupies a unique place in the theory of the calculus of variations in the large—a place which invites further study—it is not to this application that this investigation is directed. My principal objective is to study the relationships between the categories and the standard topological invariants (homology and homotopy groups, homotopy type, etc.). This is a particularly interesting view-point, first noted by Borsuk [14], because the categories seem to have a large measure of independence from these invariants.

The carrying out of this program is greatly facilitated by "dimensionalization" of the categories. It is reasonable to expect that such an analysis of the categories should prove to be a useful tool. This dimensionalization is relatively easy to carry out in the case of homology category and has been implicitly indicated by Eilenberg [20, p. 187]; in the case of the homotopy category it is necessary to define an "n-dimensional homotopy." It is hoped that this dimensionalized homotopy will be of much wider use than in the study of category. There is a strong analogy between the homotopy categories and the homotopy groups, and between the homology categories and the homology groups.

In Chapter 1, I redefine the category, making use of open, instead of the previously used closed, coverings. This is by no means a trivial change. The old definition, as Borsuk [14] has observed, applies to any topological space, while the new definition is not so generally applicable. A little reflection will convince one that those places for which the new definition fails are of little interest in connection with category. Furthermore, the new definition is readily susceptible to the dimensionalization mentioned above. Chapter 1 is written in such a way that the results are valid for the categories defined in Chapters 2 and 3.

In Chapter 2, I define *homotopy in dimension n* and introduce the corresponding *n-dimensional category*. Relations between category, *n*-dimensional category,

[1] The numbers in brackets refer to the bibliography at the end of the paper.

and various classical invariants are developed. Chapter 3 is devoted to the homology categories.

In Chapter 4, I consider invariants which I call the strong categories. The strong homotopy category was defined implicitly by Borsuk [15]. Just as in the case of the categories, a strong category is defined for each of the relations:

homology in dimension n, homology

homotopy in dimension n, homotopy.

The strong categories exhibit a much greater degree of independence than the categories; for instance, they are not dependent on either the categories, the homotopy type, or deformation retraction. Whether these independencies hold over the class of manifolds is an open and vital question.

Finally, in Chapter 5, I indicate some extensions of the notion of category; in particular a connection between the categories and the multicoherence.

In addition to new results on category many of the previous results have been extended, deepened, and recast.

I have intended this paper to be definitive; I believe that almost all previous results (except applications to calculus of variations and differential geometry) have been included. The bibliography is believed to be complete. Professors Lefschetz and Hurewicz have made many kind suggestions and criticisms. Their inspiration has been invaluable.

I. Homotopy Category

1. Definitions. Let X and M be separable, metric spaces[2] and let f_0 and f_1 be mappings[3] ϵM^X. The mappings f_0 and f_1 are said to be *homotopic in* M, written $f_0 \simeq f_1$ in M, if there is a mapping $f \epsilon M^{X \times [0,1]}$ such that $f(x, 0) = f_0(x)$ and $f(x, 1) = f_1(x)$ for every $x \epsilon X$. Such a mapping f is called a *homotopy between* f_0 and f_1. In general it is important to know in what space a given homotopy occurs, because a mapping space K^X is considered (in an obvious way) as a subspace of M^X whenever $K \subset M$. If there is no danger of confusion the phrase "in M" may be omitted.

When $X \subset M$, X is said to be *deformable in M into* $Y \subset M$ if there is a mapping $f_0 \epsilon Y^X$ which is homotopic in M to $f_1 = 1$, the identity[4] mapping of X. A homotopy between f_0 and 1 is called a *deformation of* X *into* Y.

If there is a point $m \epsilon M$ such that X can be deformed in M into m then the set X is said to be *contractible in* M. Thus X is contractible in M when the identity mapping f_1 of M^X is homotopic to a constant[5] mapping f_0 of M^X. A deformation of X into m is called a *contraction of* X *in* M.

A subset A of M will be called a *categorical subset* of M if there is an open

[2] In this paper only separable metric spaces will be considered.

[3] By a *mapping* of M^X is meant a continuous function defined on X with values in M.

[4] An identity mapping of X is the function defined by $f(x) = x$ for every $x \epsilon X$.

[5] A constant mapping f_0 is defined by $f_0(x) = m$ for every $x \epsilon X$. It is sometimes convenient to confuse the mapping f_0 and the point $m \epsilon M$.

set U of M which contains A and is contractible in M. Since any subset of a set contractible in M is itself contractible in M every categorical subset of M is contractible in M; the converse is not true. A covering of X by categorical subsets of M will be called a *categorical covering of X in M*. We shall denote the collection of categorical coverings of X in M by $C_M(X)$.

For any covering σ of some space let us denote by $|\sigma|$ the number of sets of σ. The *category*, $\mathrm{cat}_M X$, of X in M is here[6] defined to be the smallest of the cardinal numbers $|\sigma|$ as σ ranges over $C_M(X)$. A categorical covering σ of X in M will be said to be *minimal* if $|\sigma| = \mathrm{cat}_M X$.

2. The Basis for Abstraction. Many theorems on category do not utilize fully the special character of the relation of homotopy. In the present chapter results of this type will be developed. For this purpose I shall write $\leftrightarrow$ for any relation of equivalence (i.e. one which is symmetric, reflexive and transitive) which has the following properties:

(2.1). If $f_0 \leftrightarrow f_1$ in K then $f_0 \leftrightarrow f_1$ in M for every M which $\supset K$.

(2.2) If ϕ_0 and $\phi_1 \in X^P$ and f_0 and $f_1 \in M^X$ are such that $\phi_0 \leftrightarrow \phi_1$ and $f_0 \leftrightarrow f_1$, then $f_0\phi_0 \leftrightarrow f_1\phi_1$.

(2.3) If A and B are mutually separated,[7] M arcwise connected and f_0 and $f_1 \in M^{A+B}$ such that $f_0 | A \leftrightarrow f_1 | A$ and $f_0 | B \leftrightarrow f_1 | B$ then $f_0 \leftrightarrow f_1$.

(2.4) If f_0 and $f_1 \in M^X$ such that $f_0 \leftrightarrow f_1$ and if x_0 and $x_1 \in X$ are such that $f_0(x_0)$ and $f_1(x_0)$ belong to an arc of M then $f_0(x_1)$ and $f_1(x_1)$ also belong to an arc of M.

(2.5) If ϕ_i and $\psi_i \in M_i^{X_i}$ $(i = 1, 2)$ such that $\phi_1 \leftrightarrow \psi_1$ and $\phi_2 \leftrightarrow \psi_2$ then $\phi \leftrightarrow \psi$ where $\phi = (\phi_1, \phi_2)$ and $\psi = (\psi_1, \psi_2) \in M^X$ and $X = X_1 \times X_2$ and $M = M_1 \times M_2$.

The relation h (homotopy) has these properties. Later we shall meet relations h_n (homotopy in dimension n), H (homology) and H_n (homology in dimension n) which also have these properties. In the remainder of this chapter (till §13) the definitions in §1 of deformation, contraction and category are to be read in the sense that the underlying equivalence relation is any predetermined $\leftrightarrow$. In particular, $\leftrightarrow$ may be h, h_n, H or H_n.

Two consequences of (2.2) are sufficiently important to be explicitly pointed out:

(2.6) If f_0 and $f_1 \in M^X$ such that $f_0 \leftrightarrow f_1$ and $A \subset X$ then $f_0 | A \leftrightarrow f_1 | A$.

(2.7) The relation $\leftrightarrow$ is a topological invariant.

To demonstrate (2.6) choose $P = A \subset X$ and $\phi_0 = \phi_1 =$ identity mapping of A. Then $f_0\phi_0 = f_0 | A$ and $f_1\phi_1 = f_1 | A$. The proof of (2.7) is trivial.

---

[6] This definition is applicable only to spaces X, M which have the property that every point of X is a categorical subset of X in M. This restricts the range of the definition from the generality of the definition used by Borsuk [14]. Moreover, even when they both apply, they need not be the same. However, as we shall see later, the definitions coincide in the important case that M is an absolute neighborhood retract, (see §14).

[7] Spaces X and Y are said to be mutually separated if they are disjoint and open in their union $X + Y$.

From now on we assume that $C(M)$ is not vacuous; this condition is satisfied, for instance, when M is locally contractible.

3. Elementary Relations. The function $\text{cat}_M X$ is an increasing function of X and a decreasing function of M in the following sense:

(3.1) *If* $X \subset Y$ *then* $\text{cat}_M X \leqq \text{cat}_M Y$. *If M is open in N then* $\text{cat}_M X \geqq \text{cat}_N X$.

The proof is obvious.

Let σ and σ' be two coverings. I shall say that σ is a *precise refinement* of σ' if there is a (1, 1) correspondence between the sets of σ and σ' such that every set of σ is contained in the corresponding set of σ'.

(3.2) *There is always a minimal covering belonging to* $C_M(X)$ *which is open. If X is closed and* $\text{cat}_M X$ *is finite there is a closed minimal covering* $\epsilon\, C_M(X)$. *If X is closed and M is a complex*[8] *there is a closed minimal covering* $\epsilon\, C_M(X)$ *whose sets are subcomplex of M in a certain subdivision.*[9]

For every categorical covering of X in M is a precise refinement of an open categorical covering. Every finite open categorical covering has a closed categorical precise refinement. If M is a complex it may be subdivided so fine that every simplex is contained in at least one of the sets of a preassigned open categorical covering.

It is easy to verify that

(3.3) *If X is compact,* $\text{cat}_M X$ *is finite.*

4. Further Elementary Relations. There is a "triangle inequality" for category:

(4.1) *For any collection* $\{X_\alpha\}$ *of subsets of M,* $\text{cat}_M \left(\sum X_\alpha\right) \leqq \sum \text{cat}_M X_\alpha$.

If, for each α the covering σ_α belongs to $C_M(X_\alpha)$ then the covering σ of $\sum X_\alpha$ which consists of all the sets of all the σ_α's is a covering belonging to $C_M(\sum X_\alpha)$. Clearly, $|\sigma| = \sum |\sigma_\alpha|$.

(4.2) *If M is arcwise connected and X and Y are mutually separated*[7] *then* $\text{cat}_M (X + Y) = \max \{\text{cat}_M X, \text{cat}_M Y\}$.

Since M is completely normal, X and Y are contained in disjoint open sets

---

[8] In this paper complexes are understood to be finite.

[9] But there is no integer k such that there is always a closed minimal covering $\epsilon\, C_M(X)$ whose sets are subcomplexes of M in the k^{th} subdivision. Let $X = M$ be the 2-dimensional Möbius strip mod m, M_m^2, [2, chapters IV, V, VI, Anhang, 12] obtained from a triangulated pseudoprojective space, P_m^2, by removing the interior of a 2-simplex, Q. The category of M is clearly 2. The required number, k, of subdivisions of M must be so large that each 1-simplex of the boundary of Q is subdivided into at least $2m/3$ 1-simplexes. Thus 2^k must be $\geqq 2m/3$ so that k must be $\geqq \log_2 2/3 + \log_2 m$.

This example, kindly shown to me by S. Eilenberg, was constructed by Eilenberg and J. H. C. Whitehead, to answer the following question of H. Hopf: Can one find, for every integer j, a multicoherent complex which is "simplicially" unicoherent in the j^{th} subdivision? The complex M_m^2 has this property with j the largest integer in $\log_2 2/3 + \log_2 m$.

It would be interesting to investigate these problems further; especially with the added restriction that the complex M be a manifold.

U and V respectively. Let σ be a categorical covering of X in M by open sets of U and σ' a categorical covering of Y in M by open sets of V. The covering σ'' of X and Y whose sets are unions of pairs of sets, one from σ and one from σ', is open and, by (2.3), belongs to $C_M(X + Y)$. But σ'' contains a subcovering σ''' for which $| \sigma''' | = \max \{| \sigma |, | \sigma' |\}$. Thus $\text{cat}_M (X + Y) \leqq \max \{\text{cat}_M X, \text{cat}_M Y\}$. But, by (3.1), $\text{cat}_M (X + Y) \geqq \max \{\text{cat}_M X, \text{cat}_M Y\}$.

A space M will be said to be a *divisor* of a containing space N if for any space X and mappings f and $g \in M^X$ the homotopy of f and g in N implies their homotopy in M. For example a retract [6] M of N is a divisor of N, by (2.2).

(4.3) *If M is a divisor of N then* $\text{cat}_M X \leqq \text{cat}_N X$.

For a categorical subset of X in N, which is $\subset$ M is then a categorical subset of X in M.

(4.4) *If M_1 and M_2 are mutually separated and $X \subset M_1$ then* $\text{cat}_{M_1+M_2} X = \text{cat}_{M_1} X$.

For M_1 is a retract of $M_1 + M_2$ and is open in $M_1 + M_2$.

(4.5) *If $X = X_1 + X_2$, $M = M_1 + M_2$ and $X_1 \subset M_1$, $X_2 \subset M_2$ where M_1 and M_2 are mutually separated then* $\text{cat}_M X = \text{cat}_{M_1} X_1 + \text{cat}_{M_2} X_2$.

By (2.4), every categorical subset of X in M is a categorical subset of either X_1 or X_2 in M. Hence, $\text{cat}_M X_1 + \text{cat}_M X_2 \leqq \text{cat}_M X$. But, by (4.1), $\text{cat}_M X \leqq \text{cat}_M X_1 + \text{cat}_M X_2$ so that $\text{cat}_M X = \text{cat}_M X_1 + \text{cat}_M X_2$. By (4.4), $\text{cat}_M X_1 = \text{cat}_{M_1} X_1$ and $\text{cat}_M X_2 = \text{cat}_{M_2} X_2$.

If M is locally arcwise connected, so that the components of M are open, then, by an extension of (4.5), $\text{cat}_M X = \sum \text{cat}_{M_i} (X \cdot M_i)$, the summation extended over the components M_i of M. Thus no generality is lost in the investigation of $\text{cat}_M X$ if M is supposed connected (hence, arcwise connected). If one is willing to restrict oneself to locally connected X, it follows in the same way from (4.2) that no generality is lost if X is also assumed connected (hence arcwise connected).

5. Categorical sequences. I shall call a finite sequence $\{A_1, A_2, \cdots, A_k = X\}$ of closed subsets of X a *categorical sequence* for X in M if $A_1 \subset A_2 \subset \cdots \subset A_k$, and if $A_1, A_2 - A_1, \cdots, A_k - A_{k-1}$ are categorical subsets of M. The *length* of a categorical sequence $\{A_1, A_2, \cdots, A_k\}$ is k.

THEOREM 5.1. *If M is arcwise connected and $\text{cat}_M X$ is finite then $\text{cat}_M X$ is the minimum of the lengths of the categorical sequences for X in M. Furthermore, if X is finite dimensional, a categorical sequence $\{A_1, A_2, \cdots, A_k\}$ of minimum length can be chosen so that*

$$\dim A_1 < \dim A_2 < \cdots < \dim A_k .$$

First we prove that if $\{A_1, \cdots, A_k\}$ is a categorical sequence for X in M then $\text{cat}_M X \leqq k$. This is obvious for $k = 1$; suppose that it has been proved for $k \leqq r - 1$ and let $\{A_1, A_2, \cdots, A_r\}$ be a categorical sequence for X in M. Since A_1 is, by assumption, categorical in M there is an open set X_1 containing

A_1 which is contractible in M. It is easy to verify that $\{A_2 - X_1, A_3 - X_1, \cdots, A_r - X_1\}$ is a categorical sequence for $X - X_1$ in M of length $r - 1$. By the induction hypothesis, $\text{cat}_M (X - X_1) \leqq r - 1$. Hence, by (4.1) $\text{cat}_M X \leqq r$, completing the induction.

Next we prove that there is a categorical sequence for X in M of length $\leqq \text{cat}_M X$. This is obvious for $\text{cat}_M X = 1$; suppose that it has been proved for $\text{cat}_M X \leqq r - 1$ and let $\{X_1, \cdots, X_r\}$ be an open minimal covering belonging to $C_M(X)$. Denote by F_i the set of points of X which belong to X_j for $j \leqq i$ but not to X_j for $j > i$. Thus the sets F_i are closed in X. Since F_1 and $X - X_1$ are closed disjoint sets of the normal space X, there is an open set G_1 of X such that $F_1 \subset G_1$ and $\bar{G}_1 \cdot (X - X_1) = 0$. If X is finite dimensional, G_1 can be chosen so that $\dim X \cdot (\bar{G}_1 - G_1) < \dim X$.

Suppose we have constructed $j - 1$, open sets $G_1, \cdots, G_{j-1}$ of X such that, for $i \leqq j - 1$,

(5.2) $G_i \supset F_i - (G_1 + \cdots + G_{i-1})$, and

(5.3) $\bar{G}_i \cdot (X - X_i) = 0$,

and, if X is also finite dimensional, $\dim X \cdot (\bar{G}_i - G_i) < \dim X$.

The sets $X - X_j$ and $F_j - \sum_{i<j} G_i$ are closed in X. Since $(G_1 + \cdots + G_{j-2}) + G_{j-1} \supset F_{j-1}$ we have

$$(F_j - \sum_{i<j} G_i) \cdot (X - X_j) \subset (F_j - F_{j-1}) \cdot (X - X_j) \subset X_j \cdot (X - X_j) = 0.$$

Hence there is an open set G_j of X containing $F_j - \sum_{i<j} G_i$ and such that $\bar{G}_j \cdot (X - X_j) = 0$. Thus we construct inductively r open sets $G_1, \cdots, G_r$ which satisfy (5.2) and (5.3) for every $i \leqq r$. If X is finite dimensional, $\dim X \sum_{i \leqq r} (\bar{G}_i - G_i) < \dim X$.

Since $\bar{G}_1 - G_1 \subset X_1 - F_1 \subset X_2 + \cdots + X_r$, it follows, by (5.3), that

$$\sum_{i \leqq r} (\bar{G}_i - \bar{G}_i) \subset X_2 + \cdots + X_r,$$

so that the category of $\sum_{i \leqq r} (\bar{G}_i - G_i)$ in M is $\leqq r - 1$. Hence, by the induction hypothesis, there is a categorical sequence $\{A_1, \cdots, A_{k-1}\}$ for $\sum_{i \leqq r} (\bar{G}_i - G_i)$ in M whose length $k - 1$ is $\leqq r - 1$.

Since $X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i)$ is closed in X, the sets $X \cdot A_j$ are closed in X. In order to show that $\{X \cdot A_1, \cdots, X \cdot A_{k-1}, X\}$ is a categorical sequence for X in M it remains only to show that $X - X \cdot A_{k-1}$ is categorical in M. But $X - X \cdot A_{k-1} = X - X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i)$ is open in X and every component is contained in one of the sets $G_i \subset X_i$. Since each X_i is contractible in M it follows from (3.1) and (4.2) that $X - X \cdot A_{k-1}$ is categorical in M.

If X is finite dimensional the last statement of the theorem follows inductively from the possibility of choosing the G's in such a way that $\dim X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i) < \dim X$.

From theorem 5.1 we can quickly deduce the Lusternik-Schnirelmann-Borsuk theorem, [14, Satz 4; 18; 40, p. 33; 42, p. 132].

(5.4) *If M is arcwise connected and* $\text{cat}_M X$ *is finite then* $\text{cat}_M X \leqq 1 + \dim X$.

If X is not finite dimensional there is nothing to prove. For finite dimensional

X this is an immediate consequence of the existence of a categorical sequence of minimum length whose sets are of increasing dimension.

6. A property of minimal coverings. If M is arcwise connected and the *finite covering σ of X by open sets of M is a minimal categorical covering of X in M then*

a) *the nerve of σ is a simplex and the image of X under the Alexandroff mapping*[10] *intersects every open face of this simplex.*

b) *The distributive lattice generated by the sets of σ, under the operations of union and intersection, is free.*[11]

Let $\sigma = \{X_1, \cdots, X_k\}$ be an open minimal categorical covering of X in M. Since X is normal there is a closed covering $\{W_1, \cdots, W_k\}$ with $W_i \subset X_i$, (3.2). Let T_i, $i = 1, \cdots, k$, denote the set of points of X which belong to at least $k - i + 1$ of the sets $W_1, \cdots, W_k$. Using (4.2) it is not difficult to see that $\{T_2 - T_1, \cdots, T_k - T_1\}$ is a categorical sequence for $X - T_1$ in M so that $\mathrm{cat}_M(X - T_1) \leqq k - 1$. Since $\mathrm{cat}_M X = k$ by hypothesis, $T_1 \neq 0$. Hence $X_1 \cdot X_2 \cdots X_k \neq 0$, i.e. the nerve of σ is a simplex (of dimension $k - 1$).

The inverse images of the open faces of this simplex under the Alexandroff mapping are the sets $X_{i_1} \cdot X_{i_2} \cdots X_{i_j} - \sum' X_i$, $i \leqq j \leqq k$, where the summation extends over $i \neq i_1, i_2, \cdots, i_j$. It is to be shown that no such set is vacuous. Suppose, for example, that $X_1 \cdots X_j \subset X_{j+1} + \cdots + X_k$. Choose a closed covering $\{W_1, \cdots, W_k\}$ which is a precise refinement of σ, with $W_i \subset X_i$.

The j closed sets $W_i - (X_{j+1} + \cdots + X_k)$, $i = 1, \cdots, j$, have a vacuous intersection; hence we may choose open sets $U_1, \cdots, U_j$, satisfying

$$U_1 \cdot U_2 \cdots U_k = 0 \quad \text{and} \quad W_i - (X_{j+1} + \cdots + X_k) \subset U_i \subset X_i .$$

Let $\sigma' = \{U_1, U_2, \cdots, U_j, X_{j+1}, \cdots, X_k\}$. Then $|\sigma'| = k = \mathrm{cat}_M X$, so that σ' is a minimal covering of $C_M(X)$. On the other hand $U_1 \cdot U_2 \cdots U_k = 0$, so that the nerve is not a simplex. Hence the set $X_1 \cdots X_j - (X_{j+1} + \cdots + X_k)$ is not vacuous.

To prove the second statement it is sufficient to show that a distributive lattice generated by elements $X_1, \cdots, X_k$ is free if no relation of the form $X_1 \cdots X_j + X_{j+1} + \cdots + X_k = X_{j+1} + \cdots + X_k$ holds.[12]

---

[10] By the Alexandroff mapping I mean here the one defined, say, in [30, p. 93].

[11] For the definition of distributive lattice the reader is referred to [5]. A distributive lattice is free if the only relations are those implied by the axioms for a distributive lattice.

[12] A relation of the lattice is of the form

$$\sum_i \prod_j X_{m_{ij}} = \sum_i \prod_j X_{n_{ij}}; \qquad m_{ij} \neq m_{ik}; \qquad n_{ij} \neq n_{ik} .$$

Let $X_1 \cdots X_t$ be any product of shortest length in this relation and let $X_{t+1}, \cdots, X_k$ be the rest of the elements. Then, adding $X_{t+1} + \cdots + X_k$ to both sides, $X_1 \cdots X_t + X_{t+1} + \cdots + X_k = X_{t+1} + \cdots + X_k$, since every product, with the single exception of $X_1 \cdots X_t$, contains at least one of the elements $X_{t+1}, \cdots, X_k$.

A_1 which is contractible in M. It is easy to verify that $\{A_2 - X_1, A_3 - X_1, \cdots, A_r - X_1\}$ is a categorical sequence for $X - X_1$ in M of length $r - 1$. By the induction hypothesis, $\operatorname{cat}_M (X - X_1) \leqq r - 1$. Hence, by (4.1) $\operatorname{cat}_M X \leqq r$, completing the induction.

Next we prove that there is a categorical sequence for X in M of length $\leqq \operatorname{cat}_M X$. This is obvious for $\operatorname{cat}_M X = 1$; suppose that it has been proved for $\operatorname{cat}_M X \leqq r - 1$ and let $\{X_1, \cdots, X_r\}$ be an open minimal covering belonging to $C_M(X)$. Denote by F_i the set of points of X which belong to X_j for $j \leqq i$ but not to X_j for $j > i$. Thus the sets F_i are closed in X. Since F_1 and $X - X_1$ are closed disjoint sets of the normal space X, there is an open set G_1 of X such that $F_1 \subset G_1$ and $\bar{G}_1 \cdot (X - X_1) = 0$. If X is finite dimensional, G_1 can be chosen so that $\dim X \cdot (\bar{G}_1 - G_1) < \dim X$.

Suppose we have constructed $j - 1$, open sets $G_1, \cdots, G_{j-1}$ of X such that, for $i \leqq j - 1$,

(5.2) $G_i \supset F_i - (G_1 + \cdots + G_{i-1})$, and

(5.3) $\bar{G}_i \cdot (X - X_i) = 0$,

and, if X is also finite dimensional, $\dim X \cdot (\bar{G}_i - G_i) < \dim X$.

The sets $X - X_j$ and $F_j - \sum_{i<j} G_i$ are closed in X. Since $(G_1 + \cdots + G_{j-2}) + G_{j-1} \supset F_{j-1}$ we have

$$(F_j - \sum_{i<j} G_i) \cdot (X - X_j) \subset (F_j - F_{j-1}) \cdot (X - X_j) \subset X_j \cdot (X - X_j) = 0.$$

Hence there is an open set G_j of X containing $F_j - \sum_{i<j} G_i$ and such that $\bar{G}_j \cdot (X - X_j) = 0$. Thus we construct inductively r open sets $G_1, \cdots, G_r$ which satisfy (5.2) and (5.3) for every $i \leqq r$. If X is finite dimensional, $\dim X \sum_{i \leqq r} (\bar{G}_i - G_i) < \dim X$.

Since $\bar{G}_1 - G_1 \subset X_1 - F_1 \subset X_2 + \cdots + X_r$, it follows, by (5.3), that

$$\sum_{i \leqq r} (\bar{G}_i - \bar{G}_i) \subset X_2 + \cdots + X_r,$$

so that the category of $\sum_{i \leqq r} (\bar{G}_i - G_i)$ in M is $\leqq r - 1$. Hence, by the induction hypothesis, there is a categorical sequence $\{A_1, \cdots, A_{k-1}\}$ for $\sum_{i \leqq r} (\bar{G}_i - G_i)$ in M whose length $k - 1$ is $\leqq r - 1$.

Since $X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i)$ is closed in X, the sets $X \cdot A_j$ are closed in X. In order to show that $\{X \cdot A_1, \cdots, X \cdot A_{k-1}, X\}$ is a categorical sequence for X in M it remains only to show that $X - X \cdot A_{k-1}$ is categorical in M. But $X - X \cdot A_{k-1} = X - X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i)$ is open in X and every component is contained in one of the sets $G_i \subset X_i$. Since each X_i is contractible in M it follows from (3.1) and (4.2) that $X - X \cdot A_{k-1}$ is categorical in M.

If X is finite dimensional the last statement of the theorem follows inductively from the possibility of choosing the G's in such a way that $\dim X \cdot \sum_{i \leqq r} (\bar{G}_i - G_i) < \dim X$.

From theorem 5.1 we can quickly deduce the Lusternik-Schnirelmann-Borsuk theorem, [14, Satz 4; 18; 40, p. 33; 42, p. 132].

(5.4) *If M is arcwise connected and $\operatorname{cat}_M X$ is finite then $\operatorname{cat}_M X \leqq 1 + \dim X$.*

If X is not finite dimensional there is nothing to prove. For finite dimensional

X this is an immediate consequence of the existence of a categorical sequence of minimum length whose sets are of increasing dimension.

6. A property of minimal coverings. If M is arcwise connected and the *finite covering σ of X by open sets of M is a minimal categorical covering of X in M then*

a) *the nerve of σ is a simplex and the image of X under the Alexandroff mapping*[10] *intersects every open face of this simplex.*

b) *The distributive lattice generated by the sets of σ, under the operations of union and intersection, is free.*[11]

Let $\sigma = \{X_1, \cdots, X_k\}$ be an open minimal categorical covering of X in M. Since X is normal there is a closed covering $\{W_1, \cdots, W_k\}$ with $W_i \subset X_i$, (3.2). Let T_i, $i = 1, \cdots, k$, denote the set of points of X which belong to at least $k - i + 1$ of the sets $W_1, \cdots, W_k$. Using (4.2) it is not difficult to see that $\{T_2 - T_1, \cdots, T_k - T_1\}$ is a categorical sequence for $X - T_1$ in M so that $\operatorname{cat}_M (X - T_1) \leqq k - 1$. Since $\operatorname{cat}_M X = k$ by hypothesis, $T_1 \neq 0$. Hence $X_1 \cdot X_2 \cdots X_k \neq 0$, i.e. the nerve of σ is a simplex (of dimension $k - 1$).

The inverse images of the open faces of this simplex under the Alexandroff mapping are the sets $X_{i_1} \cdot X_{i_2} \cdots X_{i_j} - \sum' X_i$, $i \leqq j \leqq k$, where the summation extends over $i \neq i_1, i_2, \cdots, i_j$. It is to be shown that no such set is vacuous. Suppose, for example, that $X_1 \cdots X_j \subset X_{j+1} + \cdots + X_k$. Choose a closed covering $\{W_1, \cdots, W_k\}$ which is a precise refinement of σ, with $W_i \subset X_i$.

The j closed sets $W_i - (X_{j+1} + \cdots + X_k)$, $i = 1, \cdots, j$, have a vacuous intersection; hence we may choose open sets $U_1, \cdots, U_j$, satisfying

$$U_1 \cdot U_2 \cdots U_k = 0 \quad \text{and} \quad W_i - (X_{j+1} + \cdots + X_k) \subset U_i \subset X_i .$$

Let $\sigma' = \{U_1, U_2, \cdots, U_j, X_{j+1}, \cdots, X_k\}$. Then $|\sigma'| = k = \operatorname{cat}_M X$, so that σ' is a minimal covering of $C_M(X)$. On the other hand $U_1 \cdot U_2 \cdots U_k = 0$, so that the nerve is not a simplex. Hence the set $X_1 \cdots X_j - (X_{j+1} + \cdots + X_k)$ is not vacuous.

To prove the second statement it is sufficient to show that a distributive lattice generated by elements $X_1, \cdots, X_k$ is free if no relation of the form $X_1 \cdots X_j + X_{j+1} + \cdots + X_k = X_{j+1} + \cdots + X_k$ holds.[12]

[10] By the Alexandroff mapping I mean here the one defined, say, in [30, p. 93].

[11] For the definition of distributive lattice the reader is referred to [5]. A distributive lattice is free if the only relations are those implied by the axioms for a distributive lattice.

[12] A relation of the lattice is of the form

$$\sum_i \prod_j X_{m_{ij}} = \sum_i \prod_j X_{n_{ij}}; \qquad m_{ij} \neq m_{ik}; \qquad n_{ij} \neq n_{ik}.$$

Let $X_1 \cdots X_t$ be any product of shortest length in this relation and let $X_{t+1}, \cdots, X_k$ be the rest of the elements. Then, adding $X_{t+1} + \cdots + X_k$ to both sides, $X_1 \cdots X_t + X_{t+1} + \cdots + X_k = X_{t+1} + \cdots + X_k$, since every product, with the single exception of $X_1 \cdots X_t$, contains at least one of the elements $X_{t+1}, \cdots, X_k$.

7. Deformation. Now I shall show that deformation can not lower the category. More precisely:

THEOREM 7.1. *If X is open in M and can be deformed in M into Y then* $\mathrm{cat}_M X \leqq \mathrm{cat}_M Y$.

By hypothesis there is a mapping $f_0 \in Y^X$ such that $f_0 \leftrightarrow 1 \mid X$ in M. Let $\sigma = \{Y_i\}$ be an open contractible covering of Y in M, so that the covering $f_0^{-1}(\sigma) = \{f_0^{-1}(Y_i)\} = \{X_i\}$ is open in X and hence in M. Since $1 \mid Y_i \leftrightarrow$ constant, $1 \mid X_i \leftrightarrow f_0 \mid X_i \leftrightarrow$ constant. Thus X_i is contractible so that $f_0^{-1}(\sigma) \in C_M(X)$.

A property is said to be *inductive* [2, II, Anhang, 1] if, whenever each of a decreasing sequence of compact sets has the property, their intersection also has the property.

(7.2) The *property* $\mathrm{cat}_M X = n$ *for fixed M and compact X is inductive.*

To prove this it is sufficient, in view of (3.1), to show that *if $\{X_i\}$ is a sequence of subsets of M such that the closure of $\sum X_i$ is compact then there is an integer i_0 such that $\mathrm{cat}_M X_i \leqq \mathrm{cat}_M X$ for every $i \geqq i_0$, where*[13] $X = \overline{\lim}\, \{X_i\}$.

Since the closure of $\sum X_i$ is compact, X is not vacuous. There is an open set U ($= \sum X_j$ for some open minimal covering $\in C_M(X)$) containing X such that $\mathrm{cat}_M U = \mathrm{cat}_M X$. Since the closure of $\sum X_i$ is compact, U contains almost all X_i. Thus there is an integer i_0 such that $X_i \subset U$ when $i \geqq i_0$. Hence, by (3.1), $\mathrm{cat}_M X_i \leqq \mathrm{cat}_M U$.

A set $A \subset M$ will be said to be *essential* in M if no neighborhood of A can be deformed in M into a proper closed subset of A.

From (7.2) quickly follows:

(7.3) *If X is compact, there is a closed subset A of X which is essential in M and whose category in M* $= \mathrm{cat}_M X$.

By (7.2) and the irreducibility principle [2, II, Anhang, 1] there is a closed subset A of X such that $\mathrm{cat}_M A = \mathrm{cat}_M X$, but $\mathrm{cat}_M B < \mathrm{cat}_M X$ for every proper closed subset B of A. This set A is essential in M, for, if it were not, a neighborhood U could be deformed in M into a proper closed subset B of A. By (3.1) and theorem 7.1 it would follow that $\mathrm{cat}_M A \leqq \mathrm{cat}_M U \leqq \mathrm{cat}_M B$ which is in contradiction with the construction of A, since B is compact.

A closed set X will be said to be *essential in M in dimension n* if X has a closed n-dimensional subset which is essential in M; X will be said to be *essential in M in exactly $r + 1$ dimensions* if there is a sequence of integers $0 = n_0 < \cdots < n_r$ such that X is essential in M in the dimensions $n_0, \cdots, n_r$ and only these.

We can now prove the following refinement of (5.4):

THEOREM 7.4. *If M is arcwise connected and the compact set X is essential in M in exactly $r + 1$ dimensions then* $\mathrm{cat}_M X \leqq r + 1$.

The statement is true for $r = 0$ by (7.3), (3.3), and (5.4).

Suppose the theorem has been proved for $r < m$ and consider a compact X

[13] A point $x \in M$ belongs to $\overline{\lim}\, \{X_i\}$ if every neighborhood of x intersects an infinite number of the sets X_i.

which is essential in M in exactly $m + 1$ dimensions, $0 = n_0 < \cdots < n_m$. By (7.3) there is a compact set $A \subset X$ which is essential in M and such that $\text{cat}_M A = \text{cat}_M X$. Since X is inessential in M in the dimensions $> n_m$, $\dim A = k \leqq n_m$. Since A is finite dimensional there is a categorical sequence $\{A_1, \cdots, A_k\}$ for A in M of length k such that $\dim A_1 < \cdots < \dim A_k$ (theorem 5.1). Since $\dim A_{k-1} < \dim A \leqq n_m$, the compactum A_{k-1} is essential in at most m dimensions. Hence, by the induction hypothesis, $\text{cat}_M A_{k-1} \leqq m$. By (4.1), we have $\text{cat}_M A \leqq \text{cat}_M A_{k-1} + \text{cat}_M (A_k - A_{k-1}) \leqq m + 1$, completing the induction.

Further discussion of this refinement of (5.4) will follow in §18.

8. Absolute category. Of particular interest is the (absolute) category cat $M = \text{cat}_M$ of M. From (3.1) and (4.1) follows

(8.1) *If M and N are open in $M + N$ then* $\text{cat}(M + N) \leqq \text{cat } M + \text{cat } N$.

From (3.1) and (7.1) follows

(8.2) *If M is open in N, and N can be deformed in itself into M, then* $\text{cat } M \geqq \text{cat } N$.

From (3.1) and (4.3) follows

(8.3) *If M is a divisor of N then* $\text{cat } M \leqq \text{cat } N$.

For simplicity of statement I shall restrict myself in the three succeeding theorems to absolute category. Removal of this restriction is not difficult.

9. Product spaces. For the category of a product space the following inequality has been proved by Bassi [4]:

THEOREM 9. *If $M = M_1 \times M_2$ is arcwise connected and* $\text{cat } M_1$ *and* $\text{cat } M_2$ *are finite, then*

$$\max \{\text{cat } M_1, \text{cat } M_2\} \leqq \text{cat } M \leqq \text{cat } M_1 + \text{cat } M_2 - 1$$

The first inequality is an immediate consequence of (8.3) since $M_1 \times x_2$ and $x_1 \times M_2$ where (x_1, x_2), denotes a point of M, are retracts of M and are homeomorphs of M_1 and M_2 respectively.

To prove the second inequality let $\text{cat } M_1 = m$ and $\text{cat } M_2 = n$. There exists a categorical sequence $\{A_1, \cdots, A_m\}$ for M_1 in M_1 of length m; likewise a categorical sequence $\{B_1, \cdots, B_n\}$ for M_2 can be found. Suppose $m \leqq n$ and define

$$C_k = \sum_{i+j=k+1} A_i \times B_j; \quad k = 1, \cdots, m + n - 1.$$

It remains only to show that the closed sets $C_1, \cdots, C_{m+n-1} = M$ form a categorical sequence for M. One has only to verify that the sets $C_{k+1} - C_k$, $k = 1, \cdots, m + n - 2$ are categorical. But, writing $A_0 = 0$, $B_0 = 0$ for convenience,

$$C_{k+1} - C_k = \sum_{i+j=k+2} (A_i - A_{i-1}) \times (B_j - B_{j-1}), \qquad k = 1, \cdots, m + n - 2.$$

From (2.5) it follows that $(A_i - A_{i-1}) \times (B_j - B_{j-1})$ is a categorical subset of M. Furthermore, if $i < i'$, $j > j'$,

$$\begin{aligned}\overline{[(A_i - A_{i-1}) \times (B_j - B_{j-1})]}\cdot[(A_{i'} - A_{i'-1}) \times (B_{j'} - B_{j'-1})] \\ \subset \overline{(A_i - A_{i-1})}\cdot(A_{i'} - A_{i'-1}) \times M_2 \\ \subset A_i\cdot(A_{i'} - A_i) \times M_2 = 0\end{aligned}$$

and symmetrically

$$\begin{aligned}[(A_i - A_{i-1}) \times (B_j - B_{j-1})]\cdot\overline{[(A_{j'} - A_{j'-1})(B_{j'} - B_{j'-1})]} \\ \subset M_1 \times (B_j - B_{j'})\cdot B_{j'} = 0.\end{aligned}$$

Thus $C_{k+1} - C_k$ is the union of mutually separated categorical subsets. Hence, by theorem 4.2, $C_{k+1} - C_k$ is a categorical set.

10. Homotopy type. Two arcwise connected spaces, M and N, are said to have the same homotopy type [27] if there are mappings $f \in N^M$ and $g \in M^N$ such that $gf \in M^M$ is homotopic to the identity and $fg \in N^N$ is homotopic to the identity.

(10.1) *If there are mappings $f \in N^M$ and $g \in M^N$ such that $gf \in M^M$ is homotopic to the identity then* cat $M \leqq$ cat N.

Let $\sigma = \{Y_i\}$ be an open contractible covering of N and let $X_1 = f^{-1}(Y_i)$ so that $f^{-1}(\sigma) = \{X_i\}$ is an open covering of M. Since $g \mid Y_i$ is homotopic to a constant, $gf \mid X_i$ is homotopic to a constant. But $gf \mid X_i$ is homotopic to the identity mapping of M^{X_i}. Hence $f^{-1}(\sigma)$ is contractible.

(10.2) *The absolute category is an invariant of the homotopy type; i.e. if X and Y have the same homotopy type* cat X = cat Y.

11. A characterization of category. It has been shown that when M is a Hausdorff space for which $C(M)$ is not vacuous the category has the following properties:

(i) If X is a point, $\text{cat}_M X = 1$

(ii) If $X \supset Y$ then $\text{cat}_M X \geqq \text{cat}_M Y$, (3.1).

(iii) $\text{cat}_M (\sum X_\alpha) \leqq \sum \text{cat}_M X_\alpha$, (4.1).

(iv) If X is open and can be deformed in M into Y then $\text{cat}_M X \leqq \text{cat}_M Y$, (theorem 7.1).

It will now be shown that these four properties characterize, in a certain sense, the set function $\text{cat}_M X$. [42; 40].

(11) *The set of positive-integer valued functions $\lambda(X)$, defined for subsets of M and satisfying*

(i) *if X is a point then $\lambda(X) = 1$,*

(ii) *if $X \supset Y$ then $\lambda(X) \geqq \lambda(Y)$,*

(iii) $\lambda(\sum X_\alpha) \leqq \sum \lambda(X_\alpha)$,

(iv) *if X is open and can be deformed in M into Y then $\lambda(X) \leqq \lambda(Y)$,*

is partially ordered by the rule: $\lambda_1 \leqq \lambda_2$ *when* $\lambda_1(X) \leqq \lambda_2(X)$ *for every* X; *the category,* $\mathrm{cat}_M X$, *is the largest element of this partially ordered set.*

Let $\sigma = \{X_\alpha\}$ be an open minimal contractible covering of X in M. By (iv) and (i), $\lambda(X_\alpha) = 1$ for each α. By (ii) and (iii) $\lambda(X) \leqq \lambda(\sum X_\alpha) \leqq \sum \lambda(X_\alpha) = |\sigma| = \mathrm{cat}_M X$.

12. Minima of set functions. I conclude this chapter with a theorem which may be considered as the topological part of the Lusternik-Schnirelmann theorem on category and calculus of variations [38; 39; 40]. Let g be a real valued set function defined for the subsets of M and satisfying

(i*) if $X \supset Y$ then $g(X) \geqq g(Y)$.

Denote by $\mathfrak{M}^n$ the collection of sets X for which $\mathrm{cat}_M X \geqq n$ and let $c_n = \inf_{X \in \mathfrak{M}^n} g(X)$. A set X of $\mathfrak{M}^n$ will be said to be minimal (relative to $\mathfrak{M}^n$ and g) if $g(X) = c_n$.

THEOREM 12. *If* $c_m = c_n = c$ $(m < n)$, *if there exists at least one closed minimal set (relative to* $\mathfrak{M}^n$*), and if* D *is a closed set whose category in* M *is* $\leqq n - m$ *then there exists a closed minimal set (relative to* $\mathfrak{M}^m$*) disjoint to* D.[14]

There is an open set $U \supset D$ such that $\mathrm{cat}_M U = \mathrm{cat}_M D$. By assumption there is a closed minimal set X relative to $\mathfrak{M}^n$. The closed set $Y = X\cdot(M - U)$ is, by construction, disjoint to D. It remains only to show that Y is a minimal set relative to $\mathfrak{M}^n$.

Since $X \in \mathfrak{M}^n$ the category of X in M is $\geqq n$. But $X \subset Y + U$ so that, by (ii) and (iii) of §11, $\mathrm{cat}_M X \leqq \mathrm{cat}_M (Y + U) \leqq \mathrm{cat}_M Y + \mathrm{cat}_M U \leqq \mathrm{cat}_M Y + (n - m)$. Hence $\mathrm{cat}_M Y \geqq m$ so that Y belongs to $\mathfrak{M}^m$.

Since Y belongs to $\mathfrak{M}^m$ it follows that $g(Y) \geqq c$. But, since $Y \subset X$, it follows from (i*) that $g(Y) \leqq g(X) = c$. Therefore $g(Y) = c$.

Since Y belongs to $\mathfrak{M}^m$ and $g(Y) = c_m$, it is a minimal set relative to $\mathfrak{M}^m$.

II. THE n-DIMENSIONAL CATEGORY

13. Homotopy in dimension n. Among the absolute neighborhood retracts [7] the contractible spaces are characterized [25] by the vanishing of all their homotopy groups. This suggests the possibility of characterizing in an analo-

[14] The proof of this theorem uses only properties (ii) and (iii) of §11 and the existence, for any $X \subset M$, of an open neighborhood of the same category; it is therefore valid if cat_M is replaced by a positive-integer valued set function having these properties. The application of this theorem to the calculus of variations is the following: Let M be a compact, connected, finite dimensional Reimannian manifold, f a function on M of class C''. Let $g(X) = \sup_{x \in X} f(x)$ and let D be the set of points x of M where all the partial derivatives of f vanish, i.e. D is the set of stationary points of f. A theorem of Lusternik and Schnirelmann [36; 38; 40, p. 22] states that D intersects every closed minimal set. It follows that if $c_m = c_n$, and if there is a closed minimal set relative to $\mathfrak{M}^n$ and g, then $\mathrm{cat}_M D > n - m$. From this and the above mentioned theorem of Lusternik and Schnirelmann it follows, in particular, that every function of class C'' on M has at least cat M stationary points.

gous way the subsets of an absolute neighborhood retract which are contractible in it. On investigation we find the rather surprising result:

In order that a subcomplex A of a complex M be contractible in M it is not sufficient that every continuous sphere $f \epsilon M^{S_n}$, for which $f(S_n) \subset A$, be homotopic in M to a constant.

Let A be a 2-dimensional torus and M a complex obtained from A by addition of two 2-cells which span a meridian and an equator respectively, there being no other identifications. Every continuous n-sphere $f \epsilon M^{S_n}$, for which $f(S_n) \subset A$, is homotopic in M to a constant. This is clear for $n = 1$ because M has a vanishing fundamental group. For $n > 1$ it is a consequence of the fact that A is an aspherical [28] space. However, A is not contractible in M since there is a 2-cycle in A which does not bound in M.

Thus the contractibility of a subset A of an absolute neighborhood retract M can not be determined by continuous spheres alone. We shall see that a characterization may be given in terms of continuous complexes.

However, contractibility is homotopy of a very special kind. The characterization of contractibility by means of continuous complexes may be extended to a characterization of homotopy. This leads to the important notion of homotopy in dimension n:

Mappings ϕ and ψ of M^X will be said to be *homotopic in dimension n* or *n-homotopic* if for every continuous n-dimensional complex $f \epsilon X^P$ the continuous complexes ϕf and $\psi f \epsilon M^P$ are homotopic. If X is of uniform class LC^n then a necessary and sufficient condition for ϕ and ψ to be n-homotopic is that $\phi f \simeq \psi f$ for every n-dimensional compactum K and mapping $f \epsilon X^K$. The proof is along the lines of [27, §4] using [31, Theorem 5]. I shall write h for homotopy and h_n for n-homotopy.

Observe that homotopic mappings are homotopic in every dimension and that homotopy in dimension n implies homotopy in every dimension $\leqq n$.

The characterization of homotopy mentioned above is the following:

THEOREM 13. *Let A be a closed subset of an absolute neighborhood retract X and let ϕ and ψ be mappings of M^X. If there is a neighborhood U of A such that $\phi \mid U$ and $\psi \mid U$ are homotopic in every dimension $< 1 + \dim X$ then $\phi \mid A$ and $\psi \mid A$ are homotopic.*

For every $\epsilon > 0$ there is a continuous complex $f \epsilon X^P$ (where $\dim P \leqq n$ if X is n dimensional) and a mapping $g \epsilon P^X$ such that the mapping $fg \epsilon X^X$ is homotopic to the identity and $d(x, fg(x)) < \epsilon$ for every $x \epsilon X$, [35]. Choose $\epsilon < d(A, X - U)$ so that $fg(A) \subset U$.

By hypothesis and (2.2), the mappings $\phi f \mid f^{-1}(U)$ and $\psi f \mid f^{-1}(U)$ of $f^{-1}(U)$ into M are homotopic. Hence the mappings $\phi fg \mid A$ and $\psi fg \mid A$ of M^A are homotopic. But $\phi fg \mid A$ is homotopic to $\phi \mid A$ and $\psi fg \mid A$ is homotopic to $\psi \mid A$. Hence $\phi \mid A$ and $\psi \mid A$ are homotopic.

14. A subset A of M will be said to be *h_n-deformable in M* into B if there is a mapping $h_0 \epsilon M^A$, with $h_0(A) \subset B$, which is n-homotopic to the identity

mapping of M^A. I shall say that A is *h_n-contractible in M* if there is a point $m \in M$ such that A can be deformed in M in dimension n into m.

(14.1) *A closed subset A of an absolute neighborhood retract M is contractible in M if and only if there is a neighborhood of A which is contractible in M in every dimension $< 1 + \dim M$.*

The first part is a specialization of theorem 13. The second part follows from

(14.2) *If M is an absolute neighborhood retract in the weak sense*[15] *then a closed subset A is a categorical subset in M if and only if A is contractible in M* [14, theorem 3].

If A is contractible in M there is a point $m \in M$ and a mapping $h \in M^{A \times [0,1]}$ such that $h(x, 0) = m$ and $h(x, 1) = x$ for every $x \in A$. Define a mapping $h' \in M^Q$ where $Q = M \times [0] + A \times [0, 1] + M \times [1]$ by

$$h'(x, 0) = m \text{ for } x \in M$$

$$h'(x, t) = h(x, t) \text{ for } (x, t) \in A \times [0, 1]$$

$$h'(x, 1) = x \text{ for } x \in M$$

Since M is an absolute neighborhood retract in the weak sense and Q is a closed subset of $M \times [0, 1]$, the mapping h' may be extended [31, p. 276 remark 3] to a mapping $h'' \in M^G$ where G is a neighborhood of Q in $M \times [0, 1]$. Let U be an open neighborhood of A in M such that $U \times [0, 1] \subset G$. It is clear that $h'' \mid U \times [0, 1]$ is a contraction of U in M. Hence A is a categorical subset of M.

15. A subset A of M will be called an *h_n-categorical subset* of M if there is an open set U of M which contains A and is h_n-contractible in M. Clearly every h_n-categorical subset is h_n-contractible in M. In contrast to (14.2), without local assumptions on the closed set A, its h_n-contractibility in M does not imply that it is h_n-categorical in M, even with the strongest (non-trivial) assumptions on M. An example to keep in mind is the following: M is a circular ring obtained from the rectangle $|x| \leqq 2/\pi$, $|y| \leqq 1$ in the Cartesian plane by identifying the points $(-2/\pi, y)$ and $(2/\pi, y)$, A is the image under this identification of the closure of the curve $y = \cos 1/x$, $|x| \leqq 2/\pi$ and $n = 1$. In fact it is the existence of such examples which necessitates the introduction of the notion of categorical set.

A covering of X by h_n-categorical subsets of M will be called an *h_n-categorical covering of X in M*. We shall denote the collection of such coverings by $h_nC_M(X)$. The *n-dimensional* (homotopy) *category*, $h_n - \text{cat}_M X$, of X in M is defined to be the smallest of the cardinal numbers $|\sigma|$ as σ ranges over

[15] I call a separable metric space an absolute neighborhood retract in the weak sense if it is a retract of every separable metric containing space in which it is closed. Cf. [31, p. 270 footnote (1)].

$h_nC_M(X)$. A covering σ of $h_nC_M(X)$ will be said to be *minimal* if $|\sigma| = h.$ $\text{cat}_M X$.

The results of chapter 1 have been so worded that they apply to the n-dimensional category. One needs only to substitute homotopy in dimension n for $\leftrightarrow$ and make the implied changes in the succeeding definitions. To see that this is so, it is sufficient to demonstrate that homotopy in dimension n is symmetric, reflexive, and transitive, and has properties (2.1) $\cdots$ (2.5). With the possible exception of (2.5), these verifications are easy.

Let ϕ_1 and $\psi_1 \in M_1^{X_1}$ be homotopic in dimension n and also ϕ_2 and $\psi \in M_2^{X_2}$ be n-homotopic. It is to be shown that $\phi = (\phi_1, \phi_2)$ and $\psi = (\psi_1, \psi_2) \in M^X$, where $X = X_1 \times X_2$ and $M = M_1 \times M_2$, are homotopic in dimension n. Let $f \in X^P$ be a continuous n dimensional complex. Clearly $\phi f = (\phi_1\pi f, \phi_2\pi f)$ and $\psi f = (\psi_1\pi_1 f, \psi_2\pi_2 f)$, where π_1 and π_2 denote the projections of X into X_1 and X_2 respectively. By hypothesis $\phi_1\pi_1 f$ and $\psi_1\pi_1 f$ are homotopic, likewise $\phi_2\pi_2 f$ and $\psi_2\pi_2 f$ are homotopic. Hence, by property (2.5) of homotopy, ϕf and ψf are homotopic. This completes the verification of (2.5) for n-homotopy.

16. As analogue of (14.2) we have:

(16.1) *If M is of uniform class*[16] *LC^n then a subset A of uniform class LC^{n-1} is an h_n-categorical subset of M if and only if it is h_n-contractible in M.*

Since M is of uniform class LC^n there[17] is an $\epsilon > 0$ such that continuous n-dimensional complexes g_1, $g_2 \in M^P$ are homotopic whenever the distance between them is less than ϵ. Since A is of uniform class LC^{n-1} there is an $\eta > 0$ such that every partial realization of a continuous n-dimensional complex in A of mesh $< \eta$ can be completed in A to a full realization of mesh $< \epsilon/3$.

Let U be an $\eta/3$ neighborhood of A and let $f \in U^P$ be a continuous n-dimensional complex. Subdivide P so fine that the image of every simplex is of diameter $< \eta/3$. For each vertex p of P choose a point x of A such that $d(f(p), x) < \eta/3$. Define $f'(p) = x$ for every vertex p. Thus f' is a mapping of the 0-dimensional framework of P into A. The mesh of this partial realization is $< \eta$, hence f' can be completed to a full realization f'' of mesh $< \epsilon/3$. Thus $f'' \in A^P$ and $d(f, f'') < \dfrac{\eta}{3} + \dfrac{\eta}{3} + \dfrac{\epsilon}{3} < \epsilon$. Hence f and f'' are homotopic in M. But, by hypothesis, f'' is homotopic in M to a constant. Hence f is homotopic in M to a constant. Thus U is h_n-contractible in M.

(16.2) *If M is connected and of uniform class LC^1 and X is compact then there is a minimal covering of $h_1C_M(X)$ whose sets are continuous curves of dimension $\leqq \max\{1, \dim X\}$.*[18]

[16] A metric space (not necessarily compact) will be said to be of uniform class LC^n if for every $\epsilon > 0$ there is a $\delta > 0$ such that every continuous k-sphere ($k \leqq n$) of diameter $< \delta$ can be extended to a continuous $(k + 1)$-cell of diameter $< \epsilon$. For compact spaces, LC^n is identical with uniform LC^n.

[17] [34, theorem 1]. The assumption of compactness is unnecessary. In the statement K_p should be at most $(p + 1)$-dimensional.

[18] cf. [14, theorem 5]. A minimal covering of $h_nC_M(X)$ whose sets are continuous of class

Let σ be an open minimal covering of $h_1C_M(X)$ and let σ' be a precise closed refinement. Let U be a set of σ and A the set of σ' contained in U. There is a compact set B of class LC^0, which contains $A \cdot X$ and is contained in U, such that $\dim (B - A \cdot X) \leqq 1$, [14, Lemma 5; 44]. The set B has a finite number, k, of components, and, since M is arcwise connected, B may be enlarged to a continuous curve by the addition of $k - 1$ spanning arcs. Let C denote such a continuous curve constructed inductively by adding spanning arcs one at a time in such a way that at each stage the number of components is decreased by one. This set C is h_1-contractible in M because every continuous 1-dimensional complex in a component of B is homotopic to a constant in M and such a homotopy $h(x, t)$ can be found for which $h(x_0, t) = h(x_0, 0)$ for a preassigned x_0 of the antecedent and every $t \in [0, 1]$. By its construction, $\dim C \leqq \max \{1, \dim X\}$. By (16.1), C is an h_1-categorical subset of M. The collection of C's is the required minimal covering.

17. Category and n-dimensional category. It is quite clear that

(17.1) $h_n \operatorname{cat}_M X \leqq \operatorname{cat}_M X$ *and, if* $k \leqq n$, $h_k \operatorname{cat}_M X \leqq h_n \operatorname{cat}_M X$.

The question: "Under what conditions does equality hold?" receives a partial answer in

(17.2) *If X is a closed subset of an n-dimensional absolute neighborhood retract, M, then* $h_n \operatorname{cat}_M X = \operatorname{cat}_M X$.

Let σ be an open covering of $h_nC_M(X)$ and let $\sigma' \in h_nC_M(X)$ be a precise closed refinement (3.2). By (14.1), each set of σ' is contractible in M. Hence, by (14.2), σ' is a categorical covering of X in M so that $\operatorname{cat}_M X \leqq |\sigma'| = h_n \operatorname{cat}_M X$.

Another partial answer is given by

(17.3) *Let X be a finite dimensional closed subset of an aspherical absolute neighborhood retract M. Then* $\operatorname{cat}_M X = h_1 \operatorname{cat}_M X$.

By (16.2), there is a minimal covering σ of $h_1C_M(X)$ whose sets are finite dimensional continuous curves. Since each set of σ is h_1-contractible in the aspherical absolute neighborhood retract M, each set is contractible in M [28, Principal theorem]. Since M is an absolute neighborhood retract it follows from (14.2) that each set of σ is categorical in M. Thus $\operatorname{cat}_M X \leqq |\sigma| = h_1 \operatorname{cat}_M X$.

18. An upper bound for the category. According to Borsuk [11, p. 254] a compact space M is called a homotopy membrane in dimension k if every closed at most k-dimensional subset is contractible in M. From (14.2) it follows that a homotopy membrane in dimension k which is also an absolute neighborhood retract is inessential in all dimensions $\leqq k$. Hence, from theorem 7.4 we have

---

LC^k ($k \geqq n - 1$) can be chosen if X is compact and M is connected and of uniform class LC^n and has the following property: For every compact set A and open set $U \supset A$ there is a compact set B of class LC^k such that $A \subset B \subset U$. It is a little difficult to see what this condition means for $k > 0$.

(18.1) *If the m-dimensional absolute neighborhood retract M is a homotopy membrane in dimension* $k(\leqq m)$ *then* cat $M \leqq m - k + 1$.

Since the vanishing of the first k homotopy groups of a finite dimensional absolute neighborhood retract M implies that M is a homotopy membrane in dimension k, [25 Satz 5] and, since the first k homotopy groups of an absolute neighborhood retract M vanish if and only if M is simply connected and acyclic in the dimensions $\leqq k$, [26],

THEOREM 18.2. *If the m-dimensional absolute neighborhood retract M is simply connected and acyclic in the dimension* $\leqq k(\leqq m)$ *then* cat $M \leqq m - k + 1$.

Theorem 18.2 differs from (18.1) only if M is not a complex, for a complex is a homotopy membrane in dimension k if *and only if* it is simply connected and acyclic in the first k dimensions [26, 3′]. The simply connected absolute neighborhood retract $B(k, m)$, $2 \leqq k \leqq m$, of Borsuk [11, p. 256] is a homotopy membrane in dimension $m - 1$ but is acyclic only in the dimensions $\leqq k - 1$. Thus cat $B(k, m) \leqq m - k + 2$ by theorem 18.2 but cat $B(k, m) \leqq 2$ by (18.1).

In the bounds given by theorem 7.4, (18.1), and theorem 18.2 the equality sign need not hold, even when M is a manifold. In fact let $M = S_1 \times S_2$. According to corollary 20.3, cat $M = 3$. But M is essential in dimensions 0, 1, 2, and 3 so that cat $M \leqq 4$ is the best bound obtainable from the above mentioned theorems.

A similar upper bound for the category follows from an unpublished result of W. Hurewicz: If M is a simply connected m-dimensional complex and the last k Betti groups, with the real number mod. 1 for coefficient domain, vanish, then M can be deformed into its $(m - k)$ dimensional framework. Hence from theorem 7.1 and (5.4) we conclude:

(18.3) *If M is a simply connected m-dimensional complex and* $\beta_i(M) = 0$ *for* $i = m - k + 1, \cdots, m$ *(real numbers* mod 1) *then* cat $M \leqq m - k + 1$.

In a similar vein:

(18.4) *If M is a connected aspherical absolute neighborhood retract whose fundamental group is free with a finite number of generators then* cat $M \leqq 2$.

For, under the conditions of the theorem, there is a compact at most 1-dimensional set which is a deformation retract of M, [21]. As above, the statement follows from theorem 7.1 and (5.4).

19. We note some further results connected with the notion of homotopy type.

(19.1) *If M is a finite dimensional aspherical absolute neighborhood retract with Abelian fundamental group then* cat $M = h_1$ cat $M = 1 + b_1(M)$, *where* $b_1(M)$ *denotes the* 1*-dimensional Betti number of M.*

For, under the conditions of the theorem, M belongs to the homotopy type of the $b_1(M)$-dimensional torus [28]. The result follows from (10.2) and corollary 20.3.

(19.2) *If M is an absolute neighborhood retract and if for every* $\epsilon > 0$ *there is an* ϵ*-mapping of M into a metric space M′ then* cat $M \geqq$ cat M'.

This follows from (10.2) and a theorem of Eilenberg [22].

20. Product spaces. From theorem 9 it follows that, if $M = M_1 \times \cdots \times M_k$ is arcwise connected,

$$\max \{\text{cat } M_i\} \leqq \text{cat } M \leqq 1 + \sum_{i=1}^{k} (\text{cat } M_i - 1), \text{ and} \tag{20.1}$$

$$\max \{h_n \text{ cat } M_i\} \leqq h_n \text{ cat } M \leqq 1 + \sum_{i=1}^{k} (h_n \text{ cat } M_i - 1). \tag{20.2}$$

The simplest examples show that the lower bounds in (20.1) and (20.2) are not always attained; I shall now show that it is the same with the upper bounds in (20.1) and, for $n > 1$, in (20.2).

Let $k = 2$, $M_i =$ the three dimensional pseudoprojective space $P^3_{m_i}$, $i = 1, 2$, [2, VI Anhang 6, 7, 8, p. 266] where m_1 and m_2 are relatively prime. It is easy to show, by construction[19] of a categorical covering, that cat $P^3_{m_i} = h_2$ cat $P^3_{m_i} = 2$. The complexes $M = M_1 \times M_2$ and $T = x_1 \times M_2 + M_1 \times x_2$, where $(x_1, x_2) \epsilon M$, are simply connected. Moreover the homology groups of M and T are the same in every dimension, and for each r the natural homomorphism of $\beta_r(T)$ into $\beta_r(M)$ is an isomorphism which covers $\beta_r(M)$ [2, VII, §2, 3]. Hence, by an unpublished theorem of Hurewicz, T is a deformation retract of M. Since, obviously, h_2 cat $T =$ cat $T = 2$, it follows from theorem 7.1 that h_2 cat $M =$ cat $M = 2$. Thus the upper bound, 3 in this case, is not attained. (I do not know whether the upper bound is always attained if the further assumption that M be essential is imposed.)

A lower bound of a different type has been given for the homotopy category by Eilenberg [19].

THEOREM 20.3. *If $M = M_1 \times \cdots \times M_k$ is an absolute neighborhood retract and essential then* cat $M_i \geqq 2$ *for* $i = 1, \cdots, k$ *implies that* cat $M \geqq k + 1$.

The proof of this theorem, which I omit, depends on the following lemma on homotopy:

(20.4) *If A is a closed subset of an absolute neighborhood retract M and $h' \epsilon M^{A \times [0,1]}$ is a deformation of A in M then there is a deformation $h \epsilon M^{M \times [0,1]}$ such that* $h \mid A \times [0, 1] = h'$.

I do not know whether the theorem or the lemma retains its validity for n-homotopy.

Notice that the example above shows that the condition that M be essential in theorem 20.3 can not be removed, or even be replaced by the condition that each M_i be essential.

Examination of the proof of theorem 20.3 reveals that the condition that M be essential can be replaced by the apparently weaker condition that M can

[19] It is not difficult to verify that the category of a join AB is min {cat A, cat B}. The join A and B of two spaces A and B is the space obtained from the product $A \times B \times [0, 1]$ by identifying $(x, y, 0)$ with x and $(x, y, 1)$ with y for every $x \epsilon A$ and $y \epsilon B$. The pseudo-projective space P^3_m is the join of the 2-dimensional pseudo-projective space P^2_m and a 0-sphere S_0.

not be deformed into M^p, the set of points $x = (x_1, \cdots, x_k)$ of M which have at least one coordinate identical with the corresponding coordinate of a fixed point $p = (p_1, \cdots, p_k)$. On observing that, for $k = 2$, the category of M^p is the maximum of cat M_1 and cat M_2 (the proof is similar to the proofs of (4.2) and theorem 22.2), it follows from theorem 7.1 that

If $M = M_1 \times M_2$ is an absolute neighborhood retract then cat $M = 2$ *if and only if M can be deformed into M^p and* cat M_1 = cat $M_2 = 2$.

This is illustrated by the above example.

The lower bound in theorem 2.15 and the upper bound (20.1) coincide when the category of each component is exactly 2. Hence

COROLLARY 20.5. *If $M = M_1 \times \cdots \times M_k$ is an essential absolute neighborhood retract and if* cat $M_i = 2$ *for $i = 1, \cdots, k$ then* cat $M = k + 1$.

Illustrative of this corollary, the product of k spheres $S_{n_1}, \cdots, S_{n_k}$ of various dimensions has category $k + 1$. In particular the category of a k-dimensional torus is $k + 1$.

Anent the question raised above of the existence of an *essential* M with category $< 1 + \sum_{i=1}^{k}$ (cat $M_i - 1$), the method of Eilenberg [19] yields the following characterization:

If the absolute neighborhood retract $M = M_1 \times \cdots \times M_k$ is essential and if $\{B_1, B_2, \cdots, B_k\}$ is a covering of M by closed subsets, where cat$_M B_i \leqq$ cat $M_i - 1$; $i = 1, \cdots, k$, *(so that* cat $M < \sum_{i=1}^{k}$ cat $M_i - (k - 1)$*), then for some i,* $\pi_i \mid B_i$ maps B_i *essentially on* M_i. (π_i denotes the projection of M into M_i.)

Suppose that for each i, $\pi_i \mid B_i$ is essential on M_i so that $\pi_i \mid B_i$ is homotopic to a mapping $\phi_i' \in M_i^{B_i}$ such that $\phi_i'(B_i)$ is a proper subset of M_i. By (20.2) there is an extension $\phi_i \in M_i^M$ of ϕ_i' which is homotopic to π_i. Hence the identity mapping $(\pi_1, \cdots, \pi_k)$ of M is homotopic to $\phi = (\phi_1, \cdots, \phi_k)$. But $\phi(M)$ is a proper subset of M, which is impossible since M is essential. Hence, for some i, $\pi_i \mid B_i$ is essential on M_i.

It is conceivable that this situation occurs. The natural mapping of a k-sphere onto a k-dimensional projective space ($k \geqq 2$) is an example of an essential mapping which raises the category. (That the category of a k-dimensional projective space is $k + 1$ follows from (31.1).)

21. Covering spaces. If the connected space M is of class LC^0 and has an open covering which is contractible in M in dimension 1 then, to every subgroup ω of the fundamental group $\pi_1(M)$, there is defined a covering space M_ω of M [43 chapter 8]. Denote by p a fixed point of M so that the points of M_ω are classes $[\gamma]_\omega$ of continuous arcs with initial points p, γ and γ' belong to the same class $[\gamma]_\omega$ if $\gamma'\gamma^{-1}$ defines an element of the subgroup ω of $\pi_1(M)$.

THEOREM 21.1. *If M_ω is a covering space of M, where $C(M) \neq 0$, and X_ω denotes the set which lies over X then* cat$_{M_\omega} X_\omega \leqq$ cat$_M X$.

Since a set lying over an open set of M is open in M_ω, it is sufficient to show that the set of M_ω lying over a set A, which is contractible in M, is contractible in M_ω. Let A be contractible in M and $h \in M^{A \times [0,1]}$ be a deformation of A in

M into p, so that $h(A, 0) = p$ and $h(x, 1) = x$. For any $x \in A$ and $t \in [0, 1]$ and continuous arc γ in M from p to x, define

$$\gamma'_{\gamma,t}(s) = \begin{cases} \gamma\left(\dfrac{s}{t}\right) & \text{for } 0 \leqq s \leqq t, \\ h(x, 1 + t - s) & \text{for } t \leqq s \leqq 1, \end{cases}$$

so that $\gamma'_{\gamma,t}$ is a continuous arc from p to $h(x, t)$. Observe that $[\gamma'_{\gamma,t}]_\omega$ depends on $[\gamma]_\omega$ and not on the particular arc $\gamma \in [\gamma]_\omega$.

Define $h' \in M_\omega^{A_\omega \times [0,1]}$ where A_ω is the set lying over A, by $h'([\gamma]_\omega, t) = [\gamma'_{\gamma,t}]_\omega$. Then $h'([\gamma]_\omega, 1) = [\gamma]_\omega$ and $h'(A_\omega, 0) \subset p_\omega$ where p_ω denotes the set of points lying over p. Thus h' is a deformation of A into p_ω. But, by (5.4), p_ω is contractible in M. Hence A_ω is contractible in M.

THEOREM 21.2. *If M_ω is a covering space of M where $h_n C(M) \neq 0$ and X_ω denotes the set which lies over X then $h_n \operatorname{cat}_{M_\omega} X_\omega \leqq h_n \operatorname{cat}_M X$.*

The proof is analogous to the proof of the preceding theorem. Let A be h_n-contractible in M and let A_ω and p_ω have the same significance as above. Let $f \in M_\omega^P$, with $f(P) \subset A_\omega$, be a continuous n-dimensional complex. By hypothesis there is a mapping $h \in M^{P \times [0,1]}$ such that $h(P, 0) = p$ and $h(x, 1) = \phi f(x)$, where ϕ denotes the mapping of M_ω downward into M, so that ϕf is a continuous n-dimensional complex. For any $x \in P$ and $t \in [0, 1]$ and continuous arc γ in M from p to $\phi f(x)$ define

$$\gamma'_{\gamma,t}(s) = \begin{cases} \gamma\left(\dfrac{s}{t}\right) & \text{for } 0 \leqq s \leqq t, \\ h(x, 1 + t - s) & \text{for } t \leqq s \leqq 1, \end{cases}$$

so that $\gamma'_{\gamma,t}$ is a continuous arc from p to $h(x, t)$.

Define $h' \in M_\omega^{P \times [0,1]}$ by $h'(x, t) = [\gamma'_{\gamma,t}]_\omega$ where $[\gamma]_\omega = f(x)$. Then $h'(x, 1) = f(x)$ and $h'(P, 0) \subset p_\omega$. As we have seen that p_ω is contractible in M_ω it follows that f is homotopic to a constant. Hence A_ω is h_n-contractible in M.

In the two preceding theorems the equality need not hold. This follows, for instance, from the example: M = torus, M_ω = its universal covering space, the Euclidean plane.

22. Identifications. We next study the effect of certain identifications on the category.

THEOREM 22.1. *Let K_1 and K_2 be (closed) disjoint homeomorphic retracts of M where M is connected of uniform class LC^0 and $C(M) \neq 0$ (or $h_n C(M) \neq 0$). Let N be the space obtained from M by an identification $f \in N^M$ of the corresponding points of K_1 and K_2. Then $\operatorname{cat} M \leqq \operatorname{cat} N$ (or $h_n \operatorname{cat} M \leqq h_n \operatorname{cat} N$).*

Let $\tilde{N}$ denote the space obtained from a sequence $\{M^i\}$, $i = \cdots, -1, 0, 1, \cdots$ of copies of M by identifying the corresponding points of K_2^i and K_1^{i+1} for each i, where K_1^i and K_2^i are the sets of M^i which correspond to K_1 and K_2 respectively. It is no loss of generality to suppose that $M = M^0$, $K_1 = K_1^0$,

$K_2 = K_2^0$. Thus the mapping ϕ of $\tilde{N}$ into N defined by $\phi(x) = f(x^0)$, where x^0 is the point of M^0 which corresponds to $x \epsilon \tilde{N}$, is an extension of the mapping f. But ϕ is a covering mapping—that is to say, for every point $y \epsilon N$ there is a neighborhood U such that $\phi^{-1}(U) = \sum V_\alpha$, where each V_α is open in $\tilde{N}$ and $\phi \mid V_\alpha$ is a topological mapping of V_α on U. It follows [43, chapter 8] that $\tilde{N}$ is a covering space of N. Hence, by theorem 21.1 (or by theorem 21.2), cat $\tilde{N} \leqq$ cat N (or h_n cat $\tilde{N} \leqq h_n$ cat N). But since K_0 and K_1 are retracts of M, $M = M^0$ is a retract of $\tilde{N}$. Hence, by (8.37), cat $M \leqq$ cat $\tilde{N}$.

That the condition that K_1 and K_2 be retracts of M cannot be dropped can be seen from the following example: M is the 2-dimensional torus, K_1 a meridian of M, K_2 a simple closed curve disjoint to K_1 which can be contracted in M. K_1 is a retract of M and K_2 is not. The category of M is 3, as we have seen earlier, but the category of N is 2. In fact it is not difficult to construct a minimal covering $\epsilon\, C(N)$ with two sets, which are images under f of cylinders each deformable in M into K_1.

THEOREM 22.2. *If K_1 and K_2 of the preceding theorem are points and if* $1 <$ cat $M < \infty$ (*or if* $1 < h_n$ cat $M < \infty$) *then* cat $M =$ cat N (*or* h_n cat $M = h_n$ cat N).

Let $\{X_i\}$ be an open minimal categorical covering of M. We may assume that $K_1 \not\subset X_1$ and $K_2 \not\subset X_2$. (From the assumptions that cat $M > 1$ follows the existence of X_1 and X_2). For if, for example, $K_1 \subset \prod_{i=1}^{\text{cat}\, M} X_i$, choose a closed neighborhood A_1 of K_1 such that $f(A_1)$ is a categorical neighborhood of $K = f(K_1)$ in N and replace the covering $\{X_i\}$ by the open refinement $\{X_i\}$, where

$$\begin{cases} X_1' = X_1 - A_1, \\ X_i' = X_i, \qquad \text{for} \quad i > 1. \end{cases}$$

This operation can be performed simultaneously if necessary on both K_1 and K_2 so that we may assume that $K_1 \not\subset X_1$ and $K_2 \not\subset X_2$.

Let U_1 and U_2 be open neighborhoods of K_1 and K_2 respectively such that $f(U_i)$ and $f(U_2)$ are contractible in N, and $U_1 \cdot U_2 = U_1 \cdot X_1 = U_2 \cdot X_2 = 0$. Let $K = f(K_1 + K_2)$.

The covering $\sigma = \{f(X_i) - K, f(U_1 + U_2)\}$ of N is open and categorical in N. Furthermore $|\sigma| = 1 +$ cat M. But $(f(X_1) - K) \cdot (f(X_2) - K) \cdot f(U_1 + U_2) = 0$ so that the nerve of σ is not a simplex. Hence, by §6, σ is not a minimal covering $\epsilon\, C(N)$, so that cat $N \leqq$ cat M.

The above proof applies word for word to the n-dimensional category.

Using induction it follows from theorem 22.2 that the categories are unchanged by a succession of point identifications.

It would be worth while to generalize theorem 22.2, somehow, to the situations of theorem 22.1 (and both theorems if possible, to the type of identification $\underset{f}{+}$ considered by Borsuk [11]. An upper bound for cat N in this direction is the following (Cf. [4, p. 277. (4)] for a special case); probably too generous:

THEOREM 22.3 *Let M be an absolute neighborhood retract in the weak sense and K_1 and K_2 be disjoint homeomorphic retracts of M. Let N be the space obtained from M by an identification $f \in N^M$ of the corresponding points of K_1 and K_2. Then* cat $N \leqq$ cat $M + k$, *where* $k = \text{cat}_M K_1 = \text{cat}_M K_2$.

By (14.2), $\text{cat}_M K_1 \leqq \text{cat}\, K_1$, (for any closed covering of $C(K_1)$ is contained in an open covering of $C_M K_1)$). By (4.3), $\text{cat}\, K_1 \leqq \text{cat}_M K_1$. Thus $\text{cat}\, K_1 = \text{cat}\, K_2 = \text{cat}_M K_1 = \text{cat}_M K_2 = k$. For the same reason $\text{cat}_N K = \text{cat}\, K = k$, where $K = f(K_1) = f(K_2)$. Since $f \mid (M - (K_1 + K_2))$ is a homeomorphism, $\text{cat}_N (N - K) \leqq \text{cat}_M (M - (K_1 + K_2)) \leqq \text{cat}\, M$. Hence, by (4.1), $\text{cat}\, N \leqq \text{cat}_N (N - K) + \text{cat}_N K \leqq \text{cat}\, M + k$.

23. Category and the fundamental group. A generalization, due to Hurewicz, of a theorem of Borsuk [14] states that for the fundamental group to be free it is sufficient that the category be $\leqq 2$. If category is replaced by 1-dimensional category the condition becomes also necessary. Precisely:

THEOREM 23.1 *The 1-dimensional (homotopy) category of an LC^1-continuum M is $= 2$ if and only if the fundamental group $\pi_1(M)$ is free and non-vanishing.*

Suppose, first, that h_1 cat $M = 2$. By (16.2), there is a minimal covering $\{M_1, M_2\} \in h_1C(M)$ by open, connected sets. Since M is an LC^1 continuum, $M_1 \cdot M_2$ has a finite number of components. It follows from a theorem on the fundamental group of a union [43, chapter 7; 29] that $\pi_1(M)$ is a free group with a finite number of generators; $\pi_1(M)$ does not vanish because h_1 cat $M > 1$.

Suppose, conversely, that $\pi_1(M)$ is free and non-vanishing. Since M is an LC^1 continuum, $\pi_1(M)$ has a finite number of generators. Hence the number of generators of $\pi_1(M)$ is the coherence $r(M)$ [20, p. 175, theorem 1]. Since $\pi_1(M)$ is non-vanishing, $r(M) > 0$. Hence $M = M_1 + M_2$, where M_1 and M_2 are open and connected and $M_1 \cdot M_2$ has $1 + r(M)$ components [20, p. 172, theorem 1]. Then M_1 and M_2 must both be h_1-contractible in M, for if an element of $\pi_1(M)$ different from the identity had a representative loop in M_1, for example, then it would follow, from the above quoted theorem on the fundamental group of a union, that $\pi_1(M)$ had more than $r(M)$ generators. Hence $\{M_1, M_2\} \in h_1C(M)$ so that h_1 cat $M \leqq 2$. But h_1 cat $M \neq 1$ because $\pi_1(M)$ does not vanish.

It is really remarkable that h_1 cat $M = 2$ can be characterized by means of the fundamental group $\pi_1(M)$ alone. In fact

(23.2) *The 1-dimensional homotopy category is not an invariant of the fundamental group, even over the class of complexes.*

Let T_3 denote the 3-dimensional torus obtained from the 3-dimensional cube Q_3 of Euclidean 3-space by identifying the opposite faces in the usual way, and let W denote the 2-dimensional subcomplex of T_3 which is the image, under the identification, of the boundary of Q_3. By (17.3) and corollary 20.3, h_1 cat $T_3 = 4$, while from theorem 23.1 and (5.4) it follows that h_1 cat $W = 3$. Nevertheless T_3 and W have isomorphic fundamental groups.

From theorem 23.1 it follows that if the LC^1 continuum M is unicoherent,

its 1-dimensional category can not be 2. For if $h_1 \operatorname{cat} M = 2$ then $\pi_1(M)$ is free; if M is unicoherent it follows that $\pi_1(M) = 0$ [20, p. 175, theorem 1] thus $h_1 \operatorname{cat} M = 1$ which is a contradiction. From this follows (cf. [14]):

(23.3) *Category and n-dimensional category are not invariants of the homology groups (arbitrary modulus) even over the class of manifolds.*

The 3-sphere S_3 and a Poincare manifold L have the same homology groups (arbitrary modulus). It is clear that $h_1 \operatorname{cat} S_3 = 1$, $h_n \operatorname{cat} S_3 = 2$ for every $n > 1$ and $\operatorname{cat} S_3 = 2$. Since $\pi_1(L) \neq 0$, $h_1 \operatorname{cat} L > 1$. Since $b_1(L) \neq 0$, $\pi_1(L)$ is not free [39, chapter 7, §48] so that, by theorem 23.1, $h_1 \operatorname{cat} L \geqq 3$. Hence, by (17.1), $h_n \operatorname{cat} L \geqq 3$, for every n, and $\operatorname{cat} L \geqq 3$.

We can now calculate the categories of the compact, connected 2-dimensional manifolds. It is more or less obvious that $h_1 \operatorname{cat} S_2 = 1$, $h_n \operatorname{cat} S_2 = \operatorname{cat} S_2 = 2$ for $n > 1$. From theorem 23.1 and (5.4) it follows that for any other compact, connected, 2-dimensional manifold M, $h_n \operatorname{cat} M = M = 3$.

III. Homology Categories

24. Definitions. For any coefficient domain $\mathfrak{A}$ and positive integer k a mapping $\phi \in M^X$ induces a (natural) homomorphism of the k-dimensional homology group $\beta_k(X) = \beta_k(X, \mathfrak{A})$ of X (with coefficient domain $\mathfrak{A}$) into the k-dimensional homology group $\beta_k(M) = \beta_k(M, \mathfrak{A})$. (Homology groups are defined as Vietoris limit cycles [26, p. 521].) Mappings ϕ and $\psi \in M^X$ are said to be *homologous* ($\mathfrak{A}$) *in dimension* $n \geqq 1$, or *n-homologous* ($\mathfrak{A}$), if, for every k, $0 \leqq k \leqq n$, the same homomorphism of $\beta_k(X, \mathfrak{A})$ into $\beta_k(M, \mathfrak{A})$ is induced by ϕ and ψ. In other terms: ϕ and ψ are n-homologous ($\mathfrak{A}$) if every k-cycle ($\mathfrak{A}$), $0 \leqq k \leqq n$, of X is "mapped" by ϕ and ψ into homologous cycles of M [2, p. 211].

A subset X of M will be said to be H_n *deformable* ($\mathfrak{A}$) *in* M *into* Y if there is a mapping $f \in M^X$, with $f(X) \subset Y$, which is n-homologous ($\mathfrak{A}$) to the identity mapping of M^X. I shall say that X is H_n *contractible* ($\mathfrak{A}$) in M if there is a point $m \in M$ into which X can be H_n deformed ($\mathfrak{A}$) in M. In other words X is H_n-contractible ($\mathfrak{A}$) in M if every k cycle ($\mathfrak{A}$), $0 \leqq k \leqq n$, in X bounds in M.

I shall say that X is H_n *categorical* ($\mathfrak{A}$) in M if X is contained in an open set which is H_n contractible ($\mathfrak{A}$) in M. A covering of X by H_n categorical ($\mathfrak{A}$) subsets of M will be called an H_n *categorical* ($\mathfrak{A}$) *covering of* X *in* M; the collection of such coverings will be denoted by $H_nC_M(X) = H_nC_M(X, \mathfrak{A})$.

For future use let us observe that

(24.1) *If M is a complex, a subcomplex X is H_n categorical in M if and only if it is H_n contractible in M.*

The *n-dimensional homology category*, $H_n \operatorname{cat}_M X = H_n \operatorname{cat}(X, \mathfrak{A})$, of X in M is defined to be the smallest of the cardinal numbers $|\sigma|$ as σ ranges over $H_nC_M(X, \mathfrak{A})$. A covering σ of $H_nC_M(X, \mathfrak{A})$ is *minimal* if $|\sigma| = H_n \operatorname{cat}_M(X, \mathfrak{A})$.

Mappings ϕ and $\psi \in M^X$ are *homologous* ($\mathfrak{A}$) if they are homologous ($\mathfrak{A}$) in every dimension. As above, we define *H-deformable* ($\mathfrak{A}$) in M into B, *H-contractible* ($\mathfrak{A}$) in M, *H-categorical* ($\mathfrak{A}$) in M, and the *homology category*, $H \operatorname{cat}_M(X, \mathfrak{A})$ in terms of homology ($\mathfrak{A}$).

25. Relation to previously defined categories. The results of chapter I apply also to the homology category. In fact, homology in dimension n and homology are symmetric, reflexive, and transitive and satisfy (2.1), (2.2), and (2.3). When M is of class LC^0, (2.4) is satisfied. When X and M are absolute neighborhood retracts, (2.5) is satisfied. For then the Vietoris and singular homology groups are the same. Suppose ϕ_1 and $\psi_1 \in M_1^{X_1}$ are n-homologous and also ϕ_2 and $\psi_2 \in M_2^{X_2}$ are n-homologous. Let $X = X_1 \times X_2$, $M = M_1 \times M_2$, $\phi = (\phi_1, \phi_2)$, $\psi = (\psi_1, \psi_2)$. An n-cycle in X may be represented by a continuous complex $f \in X^P$ and a combinatorial n-cycle γ' on P. Let P' be a copy of P, τ a topological mapping of P' on P and γ' the cycle of P' corresponding to γ on P. As in the case of n-homotopy, $\phi f = (\phi_1\pi_1 f, \phi_2\pi_2 f)$ and $\psi f = (\psi_1\pi_1 f, \psi_2\pi_2 f)$. Since by hypothesis $\phi_1\pi_1$ and $\phi_2\pi_2$ are n-homologous to $\psi_1\pi_1$ and $\psi_2\pi_2$ respectively, there is a complex $Q \supset P + P'$ and a chain Γ on Q whose boundary is $\gamma - \gamma'$, and mappings $\alpha_1 \in M_1^Q$, $\alpha_2 \in M_2^Q$ such that $\alpha_1 | P = \phi_1\pi_1 f$, $\alpha_1 | P' = \psi_1\pi_1 f\tau$, $\alpha_2 | P = \phi_2\pi_2 f$, $\alpha_2 | P' = \psi_2\pi_2 f\tau$. Thus $\alpha = (\alpha_1, \alpha_2) \in M^Q$ such that $\alpha | P = \phi f$ and $\alpha | P' = \psi f\tau$. Hence ϕ and ψ are n-homologous, completing the proof that (2.5) is satisfied. In view of (24.1) it follows that §9 applies to homology, at least if M is a complex.

It is quite clear that $H_n \operatorname{cat}_M X \leqq H \operatorname{cat}_M X$, $H_k \operatorname{cat}_M X \leqq H_n \operatorname{cat}_M X$ if $k \leqq n$, $H_n \operatorname{cat}_M X \leqq h_n \operatorname{cat}_M X$ and $H \operatorname{cat}_M X \leqq \mathrm{h} \operatorname{cat}_M X (= \operatorname{cat}_M X)$. Also that $H_n \operatorname{cat}_M X = H \operatorname{cat}_M X$ when M is compact and n-dimensional.

26. Complete homology category. Mappings ϕ and $\psi \in M^X$ are said to be *completely homologous* if they are homologous ($\mathfrak{A}$) for every choice of coefficient domain $\mathfrak{A}$ [2, p. 211]. I shall say that ϕ and ψ are *completely n-homologous* if they are n-homologous ($\mathfrak{A}$) for every $\mathfrak{A}$. As in §24 we may define the *complete homology category* and the *complete n-homology category*. From (24.1) and [2, §4] of [27, Satz 1.3] it follows that the complete n-homology category of X in a complex M is $= H_n \operatorname{cat}_M (X, Z)$ where[20] Z is the direct sum of the cyclic groups Z_m of order $m \geqq 2$. If M is a complex without torsion then the complete n-homology category of X in M is $= H_n \operatorname{cat}_M (X, \mathfrak{R}_1)$ [2, §4, Nr. 13]; this statement may be false if M has torsion [2, §4, Nr. 13].

27. Covering spaces. The question naturally arises as to whether the analogue of theorems 21.1 and 21.2 holds for the homology categories. The following example shows that it does not:

Let M be the manifold with boundary obtained from a 2-dimensional torus by removing an open 2-cell and let X be the 1-sphere which is the boundary of M [2, VII, §1, Nr. 9]. The fundamental group of M is free with two generators, a and b, corresponding to an equator and a meridian of the original torus. Let M_ω be the covering space of M determined by the subgroup ω of $\pi_1(M)$ generated by a^2. Thus M_ω may be obtained from a torus by removing two open 2-cells

[20] I shall use, throughout, the following notation: Z_0 is the group of integers; Z_m is the cyclic group of order $m \geqq 2$; $\mathfrak{R}_1$ is the group of real numbers mod 1.

with disjoint closures. Clearly X_ω is the boundary of M_ω and consists of a pair of disjoint 1-spheres. It is not very difficult to see that $H_1 \operatorname{cat}_M X = 1$ while $H_1 \operatorname{cat}_M X = 2$, the coefficient domain $\mathfrak{A}$ chosen arbitrarily.

28. Homology categories and intersection cycles. A distinct advantage of the homology categories is the opportunity to apply the duality theorem to their calculation. In this number, the basis of this calculation is developed. Let $\mathfrak{B}$ be a locally compact separable group and let $\mathfrak{A}$ be the locally compact separable group of its characters [41, chapter V]. We require further that $\mathfrak{B}$ be a ring. We consider an orientable manifold,[21] M, of dimension n and a subcomplex A. A consequence of the duality theorem is:

If every r cycle ($\mathfrak{A}$) *in A bounds in M then every* $(n - r)$*-cycle* ($\mathfrak{B}$) *in M has a homologous cycle in* $M - A$, $(r > 0)$.

For there is a natural homomorphism θ of $\beta_r(M, \mathfrak{A})$ into $\beta_r(M \bmod A, \mathfrak{A})$, which reduces every r-cycle mod A. The hypothesis that every r-cycle in A bounds in M is equivalent, as an elementary argument shows, to the hypothesis that θ is an isomorphism of $\beta_r(M, \mathfrak{A})$ into $\beta_r(M \bmod A, \mathfrak{A})$. Let θ' denote the homomorphism, induced by θ, of the group $\beta_{n-r}(M - A, \mathfrak{B})$ of characters of $\beta_r(M \bmod A, \mathfrak{A})$ into the group $\beta_{n-r}(M, \mathfrak{B})$ of characters of $\beta_r(M, \mathfrak{A})$. A character λ of $\beta_r(M \bmod A, \mathfrak{A})$ is transformed by θ' into the character $\lambda\theta$ of $\beta_r(M, \mathfrak{A})$. It is easy to verify that θ' is the natural homomorphism of $\beta_{n-r}(M - A, \mathfrak{B})$ into $\beta_{n-r}(M, \mathfrak{B})$. Since θ is an isomorphism, there is, for every $\mu \in \beta_{n-r}(M, \mathfrak{B})$, a character λ' of a subgroup of $\beta_r(M \bmod A, \mathfrak{A})$ such that $\mu = \lambda'\theta$. But there is a character λ of $\beta_r(M \bmod A, \mathfrak{A})$ which is an extension of λ' [41, chapter 5, §31, theorem 35]. Since $\theta'\lambda = \lambda\theta = \mu$ we have shown that θ' covers the image group, i.e. $\theta'(\beta_{n-r}(M - A, \mathfrak{B})) = \beta_{n-r}(M, \mathfrak{B})$. This last statement means precisely that every $(n - r)$ cycle ($\mathfrak{B}$) in M has a homologue in $M - A$.

From this corollary of the duality theorem follows:

THEOREM 28.1 *If, on the orientable n-dimensional manifold M, there can be found k cycles* $\gamma_1, \cdots, \gamma_k(\mathfrak{B})$, *of dimensions* $\leqq n - 1$ *and* $\geqq n - r$, *such that their intersection cycle* [33, p. 171] $\gamma = \gamma_1 \cdot \gamma_2 \cdots \gamma_k$ *is not homologous to zero, then* $H_r \operatorname{cat}(M, \mathfrak{A}) \geqq k + 1$.

Suppose $H_r \operatorname{cat}(M, \mathfrak{A}) \leqq k$ so that, by (24.1), k subcomplexes $A_1, \cdots, A_k$ of M can be found, which cover M and are H_r contractible ($\mathfrak{A}$) in M. By the above argument, k homologues, $\delta_1, \delta_2, \cdots, \delta_k$, of $\gamma_1, \gamma_2, \cdots, \gamma_k$ respectively, can be found in $M - A_1, M - A_2, \cdots, M - A_k$ respectively. The intersection cycle $\delta = \delta_1 \cdot \delta_2 \cdots \delta_k$ is homologous to γ, so that by hypothesis the carrier $\hat{\delta}$ of δ is not empty; on the other hand $\hat{\delta} \subset (M - A_1) \cdot (M - A_2) \cdots (M - A_k) = 0$. This contradiction proves the theorem.

The corresponding theorem for the non-orientable manifolds has been proved by Schnirelmann [42; 39, p. 33; 40, p. 42] (in the case $r = n$) though not explicitly formulated:

---

[21] An n-manifold is a connected complex such that the linked complex of any r-simplex is a homology $(n - r - 1)$-sphere.

THEOREM 28.2 *If, on the n-dimensional manifold M, there can be found k cycles* $\gamma_1, \gamma_2, \cdots, \gamma_k(Z_2)$ *of dimensions* $\leqq n - 1$ *and* $\geqq n - r$, *such that their intersection cycle* $\gamma = \gamma_1 \cdot \gamma_2 \cdots \gamma_k$ *is not homologous to zero, then* $H_r \operatorname{cat}(M, Z_2) \geqq k + 1$.

The proof is along the same lines; the group Z_2 of integers mod 2 replaces both $\mathfrak{A}$ and $\mathfrak{B}$ and mod 2 characters replace the (real numbers mod 1) characters. The extension theorem corresponding to [41, theorem 35] is trivial because subgroups of homology groups (Z_2) are closed with respect to both the topology and division.

29. Homology categories and submanifolds. In the applications of §28 to the calculation of homology categories it is sometimes sufficient to use weakened forms of theorems 28.1 and 28.2. The weaker theorem (for the case Z_2 and $r = n$) is due to Schnirelmann [42; 39, p. 32; 40, p. 40]. It is based on a sequence of manifolds with properties somewhat analogous to the categorical sequence of §5.

I shall say that a sequence $M_0, M_1, \cdots, M_k = M$ of submanifolds, of dimension $n_0, n_1, \cdots, n_k$ respectively, of an n dimensional manifold M is an S-sequence $(\mathfrak{A})$ if

$$M_0 \subset M_1 \subset \cdots \subset M_k = M \text{ and}$$

$$0 \leqq n_0 < \cdots < n_k = n \text{ and}$$

for every $i = 1, 2, \cdots, k - 1$, any $n_i - n_{i-1}$ cycle $(\mathfrak{A})$ in M_i which bounds in M bounds also in M_i.

(The last condition is somewhat analogous to the "divisor" of §4.)

The application of S-sequences to the calculation of homology category depends on the lemmas immediately below, which are consequences of the duality theorems.

A cycle γ' on a submanifold M' of a manifold M will be said to have been *extended* to a cycle γ on M if the intersection cycle $\gamma \cdot \mu'$ of γ with the fundamental cycle μ' of M' is homologous on M to γ'. This definition is for an arbitrary coefficient domain if M and M' are orientable, for coefficient domain Z_2 otherwise.

Let M be an orientable n-dimensional manifold, $\mathfrak{A}$ and $\mathfrak{B}$ as in §28.

(29.1) *If M' is an orientable m-dimensional submanifold of M, with the property that every $(m - i)$-cycle $(\mathfrak{A})$ of M' which bounds in M bounds also in M', then every i-cycle $(\mathfrak{B})$ on M' can be extended to an $(n - m + i)$-cycle $(\mathfrak{B})$ on M.*

The hypothesis that every $(m - i)$-cycle $(\mathfrak{A})$ of M' which bounds in M bounds also in M' means precisely that the natural homomorphism θ of $\beta_{m-i}(M', \mathfrak{A})$ into $\beta_{m-i}(M, \mathfrak{A})$ is an isomorphism. Let γ' be an i-cycle $(\mathfrak{B})$ on M'. According to the duality theorem, γ' is a character of $\beta_{m-i}(M', \mathfrak{A})$. Hence $\gamma'\theta^{-1}$ is a character of a subgroup of $\beta_{m-i}(M, \mathfrak{A})$. There is a character γ of $\beta_{m-i}(M, \mathfrak{A})$ which is an extension of $\gamma'\theta^{-1}$ [41, chapter 5, §31, theorem 35].

The character γ is an $(n - m + i)$ cycle $(\mathfrak{B})$ on M and $\gamma \cdot \mu'$ is homologous to γ' on M.

If we allow M and M' to be non-orientable we have in an analogous fashion (cf. §28):

(29.2) *If M' is an m-dimensional submanifold of the n-dimensional manifold M, with the property that every $(m - i)$-cycle (Z_2) of M' which bounds in M bounds also in M', then every i-cycle (Z_2) on M' can be extended to an $(n - m + i)$-cycle (Z_2) on M.*

From (29.1) and (29.2) we derive, by parallel arguments, the following two theorems of which we exhibit the proof of the first only.

THEOREM 29.3 *If there is an S-sequence $(\mathfrak{A})$, $M_0, M_1, \cdots, M_k = M$, of length $k + 1$, in the orientable n-dimensional manifold M, such that the submanifolds $M_0, M_1, \cdots, M_k$ are orientable, then $H_r \operatorname{cat}(M, \mathfrak{A}) \geqq k + 1$ for any $r \geqq \max_{i=1,\cdots,k} \{n_i - n_{i-1}\}$.*

THEOREM 29.4 *If there is an S-sequence (Z_2), $M_0, M_1, \cdots, M_k = M$ of length $k + 1$, in the n-dimensional manifold M then $H_r \operatorname{cat}(M, Z_2) \geqq k + 1$ for any $r \geqq \max_{i=1,\cdots,k} \{n_i - n_{i-1}\}$.*

PROOF OF THEOREM 29.3: Let $\mu_0, \mu_1, \cdots, \mu_k$ denote fundamental cycles $(\mathfrak{B})$ of $M_0, M_1, \cdots, M_k$ respectively. Applying (29.1) to the n_{i-1} dimensional cycle μ_{i-1} of M_i, $i = 1, 2, \cdots, k$, we construct an $(n - n_i + n_{i-1})$-dimensional cycle $(\mathfrak{B})$ $\overset{*}{\mu}_{i-1}$ on M which is an extension of μ_{i-1}. It is easy to prove inductively that the n_{k-i} dimensional intersection cycle $\overset{*}{\mu}_{k-1} \cdot \overset{*}{\mu}_{k-2} \cdots \overset{*}{\mu}_{k-i}$ is homologous to μ_{k-1}, and thus that the intersection cycle $\overset{*}{\mu}_{k-1} \cdot \overset{*}{\mu}_{k-2} \cdots \overset{*}{\mu}_0$ is homologous to μ_0, hence not homologous to zero. Since the dimension of each of the cycles $\overset{*}{\mu}_{k-1}, \overset{*}{\mu}_{k-2}, \cdots, \overset{*}{\mu}_0$ is $\leqq n - 1$ and $\geqq n - r$, it follows from theorem 28.1 that $H_r \operatorname{cat}(M, \mathfrak{A}) \geqq k + 1$.

From theorem 29.3 we deduce the interesting consequence:

(29.5) *If a manifold M has dimension $\geqq 2$ and $\beta_1(M, Z_2) \neq 0$ then $H \operatorname{cat}(M, Z_2) \geqq 3$.*

For there is a simple closed curve M_1 such that the sequence $M_0 =$ point of M_1, M_1, $M_2 = M$ is an S-sequence (Z_0).

30. Product manifolds. Consider cycles $\gamma_1^j, \cdots, \gamma_{k_j}^j(\mathfrak{B})$ of dimension $\leqq n_j - 1$ and $\geqq n_j - r_j$ on the n_j-dimensional manifold M^j $(j = 1, 2)$, where $\mathfrak{B}$ is understood to be Z_2 if the manifolds are not restricted to be orientable, such that the intersection cycles $\gamma^1 = \gamma_1^1 \cdots \gamma_{k_1}^1$ and $\gamma^2 = \gamma_1^2 \cdots \gamma_{k_2}^2$ are not homologous to zero. Then the cycles $\gamma_1^1 \times \mu^2, \cdots, \gamma_{k_1}^1 \times \mu^2, \mu^1 \times \gamma_1^2, \cdots, \mu^1 \times \gamma_{k_2}^2$, $\mu^1 \times \mu^2$ on $M^1 \times M^2$, where μ^1 and μ^2 are fundamental cycles of M^1 and M^2 respectively, have an intersection not homologous to zero. In fact this intersection is $\gamma^1 \times \gamma^2$, cf. [2, VII, §3, 5]. Thus

(30.1) *If on the orientable n_j-dimensional M_j $(j = 1, 2)$, there can be found k_j cycles $\gamma_1^j, \cdots, \gamma_{k_j}^j \cdot (\mathfrak{B})$ of dimensions $\leqq n_1 - 1$ and $\geqq n_1 - r_1$ such that their intersection cycles γ^1 and γ^2 are not homologous to zero then $H_r \operatorname{cat}(M^1 \times M^2, \mathfrak{A}) \geqq$*

$k_1 + k_2 + 1$, *where* $r = r_1 + r_2$. *The same statement holds about not-necessarily-orientable manifolds if* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are replaced by* Z_2.

Similarly one can show that

(30.2) *If there is an S-sequence* ($\mathfrak{A}$) *of orientable manifolds of length* $k_j + 1$ *on the orientable manifold* M^j ($j = 1, 2$) *then* H_r cat $(M^1 \times M^2, \mathfrak{B}) \geqq k_1 + k_2 + 1$ *where* $r = r_1 + r_2$ (cf. theorem 29.3). *The same statement holds if orientability is dropped and* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are replaced by* Z_2.

Thus, in particular, the homology (Z_2) category of the product of k manifolds is $\geqq k + 1$. This is analogous to theorem 20.3.

31. The projective spaces. The most impressive application of theorem 29.4 is to the projective spaces.

(31.1) *For the n-dimensional projective space* P_n *we have* H_1 cat $(P_n, Z_2) =$ cat $P_n = n + 1$.

For there is an S-sequence (Z_2) $P_0, P_1, \cdots, P_n$ of projective spaces. It is sufficient to prove that every 1-cycle in P_i which bounds in P_{i+1} bounds in P_i ($i = 1, \cdots, n-1$). This is true for $i = 1$ because no 1-cycle of P_1 bounds in P_2. For $i > 1$, suppose γ is a 1-cycle in P_i which bounds a 2-chain Γ in P_{i+1}. Choose a point of P_{i+1} not on P_i and not on the carrier $\hat{\Gamma}$ of Γ and project Γ from this point into P_i. The projected 2-chain has γ for its boundary.

Thus H_1 cat $(P_n, Z_2) \geqq n + 1$. But H_1 cat $(P_n, Z_2) \leqq$ cat $P_n \leqq n + 1$ by §25 and (5.4).

Let us observe that cat $P_n = n + 1$ has, as an immediate consequence, the often proved[22]

THEOREM 31.2 *If the* $(n-1)$*-dimensional sphere* S_{n-1} *is covered with n closed sets then at least one of these sets contains a pair of antipodal points.*

Let $f \in P_n^{Q_n}$ be the standard identification of antipodal points of the n-dimensional cell Q_n and suppose that $\{A_1, A_2, \cdots, A_k\}$ is a closed covering of S_{n-1}, the boundary of Q_n, where no A_i contains any pair of antipodal points. Let a be the center of Q_n so that $\{\widehat{a\,A_1}, \widehat{a\,A_2}, \cdots, \widehat{a\,A_k}\}$ is a closed covering of Q_n, where $\widehat{a\,A_i}$ denotes the convex join of a and A_i. Since $\{f(\widehat{a\,A_1}), f(\widehat{a\,A_2}), \cdots, f(\widehat{a\,A_k})\}$ is a covering of P_n by closed sets which are contractible in P_n, it follows that cat $P_n \leqq k$. Hence $n + 1 \leqq k$, which proves the theorem.

32. Complexes with homology category 2. Let $\mathfrak{A}$ now denote an arbitrary coefficient domain, and let K be a connected n-dimensional complex. Suppose that H_r cat $(K, \mathfrak{A}) \leqq 2$ so that $K = K_1 + K_2$ where K_1 and K_2 are subcomplexes of K (cf. 24.1) and every i-cycle ($\mathfrak{A}$) in K_1 or in K_2 bounds in K for

---

[22] [39, p. 26, lemma 1; 8, p. 178; 16; 32; 2, XII, §4, 8, Satz XI]. Cf. also [17, 24, 23]. Another consequence is *If the n-sphere* S_n *is covered with n closed sets then at least one of these sets contains a symmetric continuum.* cf. [39, p. 26, lemma 2], where this is used as the basis of the calculation of the category of P_n.

$i = 0, 1, \cdots, r$. Under these conditions the subgroup $S_i(K)$ of $\beta_i(K)$, of these elements which can be represented as cycles $\gamma_i^1 + \gamma_i^2$ with $\hat{\gamma}_i^j \subset K_j$ $(j = 1, 2)$, consists of the identity only for $i = 0, 1, \cdots, r$ [2, VII, §2, 3]. Furthermore the subgroup $N_i(K_1 \cdot K_2)$ of $\beta_i(K_1 \cdot K_2)$, consisting of cycles which bound in either K_1 or K_2, is $= \beta_i(K_1 \cdot K_2)$ for $i = 0, 1, \cdots, r$ [2, VII, §2, 4]. (The group β_0 is meant in the sense of [2, V, §1, 5].) Hence, for every $i \leqq r$, the isomorphism [2, VII, §2, 8, Satz 5] takes the form

$$\beta_i(K, \mathfrak{A}) = \beta_{i-1}(K_1 \cdot K_2, \mathfrak{A}); \qquad i \leqq r. \tag{32.1}$$

Thus we have derived a necessary condition for $\{K_1, K_2\} \in H_r C(K, \mathfrak{A})$.

From (32.1) for $i = 1$ follows

THEOREM 32.2 *If K is a connected complex and $\beta_1(K, \mathfrak{A}) \neq 0$ and is not the direct sum of groups isomorphic with $\mathfrak{A}$ then H_1 cat $(K, \mathfrak{A}) \geqq 3$.*

For $\beta_0(K_1 \cdot K_2, \mathfrak{A})$ is the direct sum of t groups isomorphic with $\mathfrak{A}$, where $t + 1$ is the number of components of $K_1 \cdot K_2$.

Theorem 32.2 is the analogon of theorem 23.1. The proof of the converse breaks down, since we can no longer use the coherence, as in theorem 23.1. In fact the converse is false: Let K be the 2-dimensional torus and $\mathfrak{A} = Z_0$. It is well known that $\beta_1(K, Z_0) = Z_0 + Z_0$; nevertheless H_1 cat $(K, Z_0) = 3$, by theorem 29.4.

From theorems 29.3 and 32.2 follows:

(32.3) *If M is an orientable manifold of dimension $\geqq 2$ and p is a prime such that $\beta_1(M, Z_p) \neq 0$, then H cat $(M, Z_p) \geqq 3$. Thus if $\beta_1(M, Z_0) \neq 0$, the complete homology category of M is $\geqq 3$.*

For if H cat $(M, Z_p) \leqq 2$ then, by theorem 32.2, $\beta_1(M, Z_p) = Z_p + \cdots + Z_p$. There is a simple closed curve α of M which carries a non-zero element z of $\beta_1(M, Z_p)$. Since $\beta_1(\alpha, Z_p)$ is the cyclic group of order p generated by z, so that the only possible homomorphisms of $\beta_1(\alpha, Z_p)$ into $\beta_1(M, Z_p)$ other than 0 are isomorphisms, the sequence $\{M_0 = \text{point of } \alpha, M_1 = \alpha, M_2 = M\}$ is an S-sequence (Z_p). Hence by theorem 29.3, H cat $(M, Z_p) \geqq 3$. The second statement follows from §26 because $\beta_1(M, Z_0) \neq 0$ implies that $\beta_1(M, Z_p) \neq 0$ for some prime p [2, V, §4, 9, p. 235].

IV. THE STRONG CATEGORIES

33. Definitions. We have considered categories associated with each of the relations homotopy, homotopy in dimension n, homology, and homology in dimension n. These relations will be denoted by $h, h_n, H = H(\mathfrak{A}), H_n = H_n(\mathfrak{A})$. In this chapter I shall write $\mathfrak{H}$ to denote an unspecified relation $\in$ the collection $\{h, h_n, H, H_n\}$.

We shall now study a category-like invariant which I have called the strong category (abbreviated cat*). This is defined by considering coverings of X by sets which are $\mathfrak{H}$-contractible (i.e. not in M but in themselves). Roughly, this amounts to confusing the set A and the space M in which it is to be $\mathfrak{H}$-contracted. Hence, it seems advisable to demand not only that the sets be

$\mathfrak{H}$-contractible in themselves but that they also possess the properties of M which were found in the preceding sections to insure a relatively complete theory. For this purpose local $\mathfrak{H}$-contractibility seems reasonable. However, the following definition is slightly more convenient: Let $\mathfrak{H}C^*(M)$ denote the collection[23] of coverings of M by neighborhood retracts of M which are $\mathfrak{H}$-contractible. (Thus when M is locally $\mathfrak{H}$-contractible so is every set of every covering of $\mathfrak{H}C^*(M)$.) An $\mathfrak{H}$-contractible neighborhood retract of M will be called a strong $\mathfrak{H}$-categorical set of M. We define the $\mathfrak{H}$-strong category,[24] $\mathfrak{H}\,\text{cat}^*\,M$, of M to be the smallest of the cardinal numbers $|\sigma|$ as σ ranges over $\mathfrak{H}C^*(M)$. (There is almost no point in considering a strong category of X in M, as, for any decent space X, it would turn out to be independent of M.) Since the points of M constitute a covering of $\mathfrak{H}C^*(M)$, the $\mathfrak{H}$-strong category is always defined. If M is an absolute neighborhood retract, its strong h-categorical sets are absolute retracts and its strong h_n-categorical sets are those neighborhood retracts whose first n-homotopy groups vanish [25, Satz 5]. A strong $\mathfrak{H}$-categorical set need not, even under the most favorable circumstances, have an $\mathfrak{H}$-contractible neighborhood. In fact the 2-cell of Alexander [1] imbedded in the obvious way in a solid torus is a strong h_1-categorical set, but no neighborhood is h_1-contractible.

34. From the definition follow immediately

(34.1) *If M_1 and M_2 are neighborhood retracts of $M_1 + M_2$ then*

$$\mathfrak{H}\,\text{cat}^*\,(M_1 + M_2) \leqq \mathfrak{H}\,\text{cat}^*\,M_1 + \mathfrak{H}\,\text{cat}^*\,M_2\,;$$

and

(34.2) $h_n\,\text{cat}^*\,M \leqq \text{cat}^*\,M$ *and, if* $k \leqq n$, $h_k\,\text{cat}^*\,M \leqq h_n\,\text{cat}^*\,M$. The proofs are obvious.

From a theorem of Hurewicz [25, Satz 6] follows

(34.3) *If M is an n-dimensional absolute neighborhood retract then* $h_n\,\text{cat}^*\,M = \text{cat}^*M$.

Similar results are easily obtained for the strong homology categories. Note that the complete n-homology category of a complex $M = H_n\,\text{cat}^*\,(M, Z_0)$, [2, V, §4, p. 228].

(34.4) *If M is an absolute neighborhood retract, and A and B are strong $\mathfrak{H}$-categorical sets of M which have only one point in common, then $A + B$ is a strong $\mathfrak{H}$-categorical set. Consequently if A' and B' are disjoint $\mathfrak{H}$-categorical sets of M then $A' + B'$ can be enlarged to a strong $\mathfrak{H}$-categorical set $A' + \alpha + B'$ by the addition of a spanning arc α.*

Since A and B are absolute neighborhood retracts so is $A + B$ [7, p. 226]. For $\mathfrak{H} = h$ the contractibility of $A + B$ is a consequence of a theorem of Aron-

[23] That the collection $\mathfrak{H}C^*(M)$ is a much more complicated invariant than $\mathfrak{H}\,\text{cat}^*\,M$ is indicated by an example due to Borsuk [13] of a 2-dimensional complex, P, in 3-dimensional Euclidean space which is $\mathfrak{H}$-contractible but for which the least value of $|\sigma| > 1$, as σ ranges over $\mathfrak{H}C^*(P)$, is 3.

[24] I shall write $C^*(M)$ for $hC^*(M)$ and $\text{cat}^*\,M$ for $h\,\text{cat}^*\,(M)$.

zajin and Borsuk [3, p. 194]. For $\mathfrak{H} = h_n$ it follows from a theorem of Kuratowski [31, p. 277]. For $\mathfrak{H} = H$ or H_n it is more or less trivial.

REMARK: The lemma is false if M does not have suitable local properties. Example in the Cartesian plane:

$$A = \{0 \leqq x \leqq 1, y = 1\} + \{x = 0, 0 \leqq y \leqq 1\} + \sum_{n=1}^{\infty} \left\{x = \frac{1}{n}, 0 \leqq y \leqq 1\right\},$$

B the set symmetric to A with respect to the origin. Every deformation of $A + B$ leaves the origin fixed.

As a consequence of (34.4) and of the theorem of Aronzajin and Borsuk quoted above we have

(34.5) *If M is an absolute neighborhood retract and σ is a minimal covering of $\mathfrak{H}C^*(M)$, then every pair of sets of σ intersects in at least two points. For $\mathfrak{H} = h$, the intersection of any pair of sets of σ is not an absolute retract.*

35. Covering spaces. By a modification of the proof of theorem 2.19 and an extension of (34.4) can be proved

(35) *If $\tilde{M}$ is a covering space of an absolute neighborhood retract M then* cat* $\tilde{M} \leqq$ cat* M *and* h_n cat* $\tilde{M} \leqq h_n$ cat* M.

36. Upper bounds for the strong category. The analogue of (5.4), namely

THEOREM 36.1 *If M is a connected n-dimensional complex then* cat* M $\leqq n + 1$.

was proved by Borsuk [15], who observed that the theorem becomes false if M is required merely to be an n-dimensional absolute neighborhood retract.

In chapters I and II, I made successive refinements of (5.4), the ultimate refinement being theorem (18.2). I shall now show that over the class of connected complexes the analogue of theorem (18.2) does not hold.

(36.2) *There is a connected m-dimensional complex, which is simply connected and acyclic in the first k dimension, for which the strong homology (Z_0) category is $>$ the upper bound $m - k + 1$ for the category of a complex satisfying these conditions.*

In our example $m = 2$, $k = 1$ and the strong homology (Z_0) category = the strong homotopy category = 3.

Let P_m^2 denote, for $m \geqq 2$, the 2-dimensional pseudoprojective space [2, VI, Anhang 6, 7, 8, p. 266] obtained from the 2-cell $r \leqq 1$ (written in polar coordinates (r, θ)) by identifying the m-points $(1, \theta)$, $(1, \theta + 2\pi/m)$, $\cdots$, $(1, \theta + 2\pi(m - 1)/m)$) for each $0 \leqq \theta \leqq 2\pi/m$. The fundamental group $\pi_1(P_m^2)$ of P_m^2 is the cyclic group of order m whose generator is carried by the simple closed curve a_m which is the image, under the identification, of the boundary, $r = 1$, of the 2-cell.

Since $\beta_1(P_m^2, Z_0) = \pi_1(P_m^2)$ [43, chapter 7, §48] it follows from theorem 32.2 and (5.4) that H_1 cat $(P_m^2, Z_0) = 3$.

Let m and n be relatively prime integers and let $X = X_{m,n}$ denote the complex

obtained from P^2_m and P^2_n by identifying a_m and a_n and, for simplicity, write $X = P_m + P_n$. Since m and n are relatively prime it follows from a previously quoted theorem on the fundamental group of a union [43, chapter 7 §52] that X is simply connected (hence also acyclic in dimension 1). Since $\beta_2(X, Z_0) \neq 0$, X is not $H(Z_0)$-contractible, so that H cat* $(X, Z_0) \geqq 2$. Suppose $\{A, B\}$ is a covering of $HC^*(X, Z_0)$. Since H_1 cat $P_m = 3$, one of the sets $A \cdot P_m$, $B \cdot P_m$, say $A \cdot P_m$, carries a 1-cycle γ which is homologous to $q\alpha_m$, $1 \leqq q \leqq m - 1$, where α_m is the generating cycle of $\beta_1(P^2_m, Z_0)$ carried by a_m. On the other hand, since P_n is not $H(Z_0)$-contractible, $A \cdot P_n$ is a proper closed subset of P_n, so that the 1-cycle γ can not bound on $P_m + A \cdot P_n$, hence not on A. This is a contradiction since A is supposed to be $H(Z_0)$-contractible.

Thus we have shown that cat* $M \leqq m - k + 1$ is not true over the class of simply connected m-dimensional complexes M acyclic in the first k dimensions. There remains a possibility that this be true if M is further restricted to range only over pseudomanifolds (or manifolds).

In a very special case the validity of the inequality cat* $M \leqq m - k + 1$ follows from a theorem of Borsuk [12, p. 58]. The simple connectivity in this special case is a consequence of the acyclicity.

(36.3) *If the connected complex M is acyclic in dimension* 1 *and can be imbedded in 3-dimensional Euclidean space then* cat* $M \leqq 2$.

37. Identification of a pair of points. The conditions under which an analogue of theorem 22.1 can be proved are apparently much more restrictive.

(37) *Let X be an absolute neighborhood retract and Y the absolute neighborhood retract obtained from X by an identification, f, of a pair of distinct points, a_1 and a_2. If B is a strong $\mathfrak{H}$-categorical set of Y then $f^{-1}(B)$ is either a strong $\mathfrak{H}$-categorical set not containing a_1 or a_2 or is the union of disjoint strong $\mathfrak{H}$-categorical sets A_1 and A_2, $A_1 \ni a_1$, $A_2 \ni a_2$, according as B does or does not contain $b = f(a_1) = f(a_2)$. Hence, by* (34.4), $\mathfrak{H}$ cat* $X \leqq \mathfrak{H}$ cat* Y.

If B does not contain b then $f^{-1}(B)$ is homeomorphic with B, so that $f^{-1}(B)$ is an $\mathfrak{H}$-contractible absolute neighborhood retract contained in X.

If B contains b then, since B is $\mathfrak{H}$-contractible, there is, for any $x \in f^{-1}(B)$ an arc in B from $f(x)$ to b. Hence there is an arc in $f^{-1}(B)$ from x to either a_1 or a_2. Thus $f^{-1}(B)$ has at most two components. But if $f^{-1}(B)$ were connected it would contain an arc from a_1 to a_2 which is contrary to the assumption that B is $\mathfrak{H}$-contractible (cf. 2.4). Hence $f^{-1}(B) = A_1 + A_2$ where A_i is the component of $f^{-1}(B)$ which contains a_i $(i = 1, 2)$. Now $f|A_1$ and $f|A_2$ are homeomorphisms so that A_i is homeomorphic to $f(A_i)$, $i = 1, 2$. Hence each A_i is strong $\mathfrak{H}$-contractible. But $f(A_i)$ is a retract of B; in fact the mapping

$$\rho(y) = \begin{cases} y & \text{for } y \in f(A_i) \\ b & \text{for } y \in B - f(A_i) \end{cases}$$

is a retraction of B into $f(A_i)$. It follows [7, 6] that $f(A_i)$ and hence A_i is an absolute neighborhood retract, hence strong $\mathfrak{H}$-categorical.

38. Point Identification on an irreducibly closed complex. In direct contrast to theorem 22.2, point identification may raise the strong categories. We need the following lemma:

(38.1) *If M is an irreducibly closed n-dimensional complex ($n \geqq 2$) with natural domain Z_m ($m = 0, 2, 3, \cdots$) [2, VII, §1, 3, 4, 5] and A is a closed subset which is $H_{n-1}(Z_m)$-contractible in an open subset U of M then $M - U$ is contained in a component of $M - A$.*

Suppose $M - U$ intersects both D_1 and D_2 where $M - A = D_1 + D_2$ and D_1 and D_2 are disjoint open sets. Let V be a closed neighborhood of A which is $H_{n-1}(Z_m)$-contractible in U. Subdivide M so fine that every simplex which intersects A is contained in V. Let μ denote an irreducible cycle (Z_m) of M and let Γ be the n-chain which is zero on every n-simplex which meets $A + D_2$ and agrees with μ on every other n-simplex of M. Since the carrier $\hat{\gamma}$ of the boundary $\gamma = F(\Gamma)$ of Γ is a subset of V it follows from the construction of V that there is an n-chain Δ in U whose boundary is γ. Hence $\Gamma - \Delta$ is an n-cycle on M. Since $\hat{\Delta}\cdot(M - U) = 0$ while $\hat{\Gamma} \supset D_1\cdot(M - U) \neq 0$, the n-cycle $\Gamma - \Delta$ is $\neq 0$. On the other hand, since $\hat{\Gamma}\cdot D_2 = 0$ and since $\hat{\Delta}$ does not intersect the non-vacuous set $D_2\cdot(M - U)$ the n-cycle $\Gamma - \Delta$ is carried by a proper subset of M. But this is impossible because M is irreducibly closed. Hence $M - U$ is contained in a component of $M - A$.

THEOREM 38.2 *If X is an irreducibly closed n-dimensional complex ($n \geqq 2$) with natural domain Z_m ($m = 0, 2, 3, \cdots$) and Y is obtained from X by an identification, f, of three distinct points a_1, a_2, a_3, then $H_{n-1}\operatorname{cat}^*(Y, Z_m) \geqq 3$.*

Since $n \geqq 2$ and Y is obviously not acyclic in dimension 1, Y is not $H_{n-1}(Z_m)$-contractible. It is therefore sufficient to show that there does not exist a covering $\{Y_1, Y_2\}$ of Y by $H_{n-1}(Z_m)$-contractible absolute neighborhood retracts. Let $b = f(a_1) = f(a_2) = f(a_3)$ and $X_i = f^{-1}(Y_i)$. Let Y_1 be the set of the covering $\{Y_1, Y_2\}$ which contains b. By (37), $X_1 = X_1^1 + X_1^2 + X_1^3$, where X_1^1, X_1^2, X_1^3 are disjoint absolute neighborhood retracts and $a_j \in X_1^j$ for $j = 1, 2, 3$. Since $X_1 + X_2 = X$ the sets $X - X_1$ and $X - X_2$ are disjoint. By (38.1) $X - U$ is contained in a component of $X - X_2$ for every open neighborhood U of X_2. Hence $X - X_2$ is connected, and therefore is contained in one of the components of X_1, say X_1^1. It follows that the connected set $X_1^2 + X_1^3 + (X - X_1)$ is disjoint to $X - X_2$, so that $X_1^2 + X_1^3 + (X - X_1)$ is a subset of X_2. Thus we have constructed a connected subset of X_2 which contains both a_2 and a_3. This is a contradiction with (37).

39. Category and strong category. Theorem 38.2 enables us to show that category and strong category need not be the same, even for pseudomanifolds.

(39) *The $\mathfrak{H}$-category and the strong $\mathfrak{H}$-category are not identical over the class of pseudomanifolds. In fact for every integer $n \geqq 2$ there is an n-dimensional pseudomanifold J_n for which $\operatorname{cat} J_n = 2$ but $H_{n-1}\operatorname{cat}^*(J_n, \mathfrak{A}) = \operatorname{cat}^* J_n = 3$.*

The pseudomanifold J_n is obtained by identifying three distinct points of the n-sphere S_n. By theorem 22.2, $\operatorname{cat} J_n = 2$; by theorem 38.2, $H_{n-1}\operatorname{cat}^*(J_n, \mathfrak{A}) \geqq$

3, [2, V, §4, 5 and VII, §1, 7]; by actual construction of a covering of $C^*(J_n)$, $\text{cat}^* J_n \leqq 3$.

From theorem 18.2 and (36.2) we see that the example of the 2-dimensional complex $X_{m,n}$ of §36 shows that category and strong category are not identical over the class of complexes. This is an essentially weaker result than (39); however the construction of this example is interesting in this connection for its own sake, as its properties seem to depend on a different idea.

Another example of the same type is the absolute neighborhood retract B constructed by Borsuk and Mazurkiewicz [9] whose category is 2 but whose strong $H_1(\mathfrak{A})$ category is infinite. Borsuk has also constructed [10] plane curves of order 3 of arbitrarily large strong category; there is even a curve of finite order whose strong category is $> \aleph_0$ [10, p. 291].

40. Homotopy type and strong category. In contrast to (10.2), I now show that

(40) *The strong $\mathfrak{H}$-category is not an invariant of the homotopy type, even over the class of 2-dimensional pseudomanifolds.*

In fact I shall construct 2-dimensional pseudomanifolds, J and K, which belong to the same homotopy type and such that $\mathfrak{H}\,\text{cat}^* J = 3$ but $\mathfrak{H}\,\text{cat}^* K = 2$.

The pseudomanifold J will be J_2 of §39, obtained from the 2-sphere S_2 by identifying three distinct points. According to (39) $\mathfrak{H}\,\text{cat}^* J = 3$ for any of our relations $\mathfrak{H}$.

Let a_1, a_2, a_3, a_4 be four distinct points of S_2 and let K be the pseudomanifold obtained from S_2 by identifying a_1 with a_2 and a_3 with a_4. Since $\beta_1(K, \mathfrak{A})$ does not vanish $\mathfrak{H}\,\text{cat}^* K \geqq 2$. But it is easily seen by actual construction of a strong $\mathfrak{H}$-categorical covering (see figure 1) that $\mathfrak{H}\,\text{cat}^* K \leqq 2$.

In order to show that J and K belong to the same homotopy type it is convenient to imbed in 3-dimensional Cartesian space (see figure 2) as follows:

Let W_u denote the circle $(x + 2)^2 + y^2 = 4$, $z = u$ for $-2 \leqq u \leqq 2$. Let D_u denote the circle $(x + 2u)^2 + y^2 = u^2$, $z = u$ for $-2 \leqq u \leqq 0$ and the circle $(x - 2u)^2 + y^2 = u^2$, $z = u$ for $0 \leqq u \leqq 2$. Let E_u denote the circle $(x + 4(u + 1))^2 + y^2 = 4(u + 1)^2$, $z = u$ for $-2 \leqq u \leqq -1$, the point $(0, 0, u)$ for $-1 \leqq u \leqq 1$, the circle $(x - 4(u - 1))^2 + y^2 = 4(u - 1)^2$, $z = 0$ for $1 \leqq u \leqq 2$. Let T^- be a topological cylinder in the half-space $z \leqq -2$ joining the circles W_{-2} and $D_{-2} = E_{-2}$, and let T^+ be a topological cylinder in the half-space $z \geqq 2$ joining the circles W_2 and $D_2 = E_2$. It is unnecessary to further specify $T = T^- + T^+$ since the mappings f, ϕ, g_t and h_t, to be described, are identities on T.

It is clear that $J = T + \sum W_u + \sum D_u$ and $K = T + \sum W_u + \sum E_u$, (the summations extending over $-2 \leqq u \leqq 2$). In the following description the symbol $\xrightarrow{F}$ will mean that the circle on the left-hand side of the symbol is transformed into the circle on the right-hand side by a translation and an irrotational similarity.

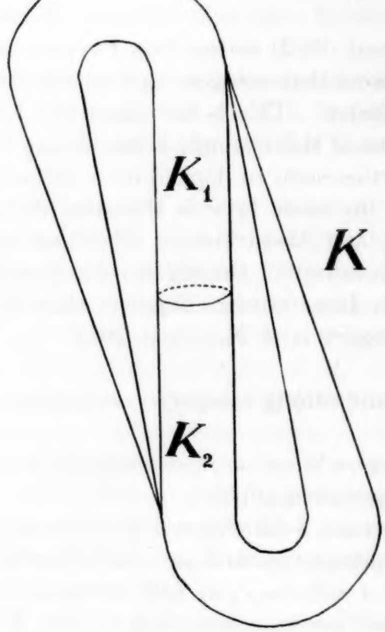


FIG. 1

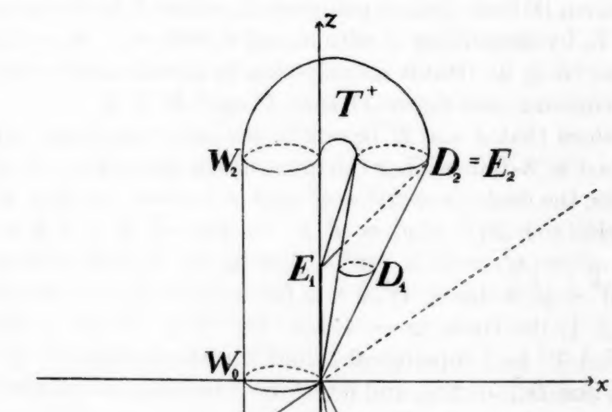


FIG. 2

The mapping $f \in K^J$ is defined by: $f \mid (T + W_u)$ is the identity, $D_u \xrightarrow{f} E_u$ for $-2 \leqq u \leqq 2$. The mapping $\phi \in J^K$ is defined by: $f \mid T$ is the identity,

$$\begin{array}{rll}
W_u \xrightarrow{\phi} W_{2(u+1)} & \text{for } -2 \leqq u \leqq -1, \\
W_0 & \text{for } -1 \leqq u \leqq 1, \\
W_{2(u-1)} & \text{for } 1 \leqq u \leqq 2, \\
E_u \xrightarrow{\phi} D_{2(u+1)} & \text{for } -2 \leqq u \leqq -1, \\
D_{2(u-1)} & \text{for } 1 \leqq u \leqq 2.
\end{array}$$

Mappings $g_t \in J^J$ and $h_t \in K^K$ are defined for every $t \in [0, 1]$ as follows:

$$\begin{array}{rll}
W_u \xrightarrow{g_t} W_{(1+t)u+2t} & \text{for } -2 \leqq u \leqq -1, \\
W_{(1-t)u} & \text{for } -1 \leqq u \leqq 1, \\
W_{(1+t)u-2t} & \text{for } 1 \leqq u \leqq 2, \\
D_u \xrightarrow{g_t} D_{(1+t)u+2t} & \text{for } -2 \leqq u \leqq -1, \\
D_{(1-t)u} & \text{for } -1 \leqq u \leqq 1, \\
D_{(1+t)u-2t} & \text{for } 1 \leqq u \leqq 2. \\
W_u \xrightarrow{h_t} W_{(1+t)u+2t} & \text{for } -2 \leqq u \leqq -1, \\
W_{(1-t)u} & \text{for } -1 \leqq u \leqq 1, \\
W_{(1+t)u-2t} & \text{for } 1 \leqq u \leqq 2, \\
E_u \xrightarrow{h_t} E_{(1+t)u+2t} & \text{for } -2 \leqq u \leqq -1, \\
E_{(1+t)u-2t} & \text{for } 1 \leqq u \leqq 2.
\end{array}$$

We observe that $g \in J^{J \times [0,1]}$ is a homotopy between $1 \in J^J$ and $\phi f \in J^J$ and that $h \in K^{K \times [0,1]}$ is a homotopy between $1 \in K^K$ and $f\phi \in K^K$. This completes the proof that the pseudomanifolds J and K belong to the same homotopy type.

Another example of the non-dependence of the strong category on the homotopy type is the following: $X_{m,n}$ is the 2-dimensional complex of §36 (m and n always relatively prime) and Y is the pseudomanifold which is the union of 2-spheres which have exactly one point in common. By (36.2) cat* $X_{m,n} = 3$; by actual construction of a covering, cat* $Y = 2$. Since $X_{m,n}$ and Y are both simply connected and $\beta_2(X_{m,n}, Z_0)$ and $\beta_2(Y, Z_0)$ are isomorphic, $X_{m,n}$ and Y belong to the same homotopy type [27, §7]. This example is weaker than the preceding one since it shows the non-dependence of the strong $\mathfrak{H}$-category only for $\mathfrak{H} = h$ and only over the class of complexes.

41. Deformation retraction and strong category. A space and a deformation retract of it belong to the same homotopy type. It will now be shown that *the strong $\mathfrak{H}$-category is not an invariant of deformation retract over the class of*

complexes. This result which is partly weaker and partly stronger than (40) implies in particular that the analogue of (8.2) is false for strong category.

Let J be as in §40 and let $L = J + \sum_{u=-1}^{1} I_u$ where I_u denotes the convex closure of D_u, $-1 \leqq u \leqq 1$. It is trivial that J is a deformation retract of L. It remains only to show that $\mathfrak{H}$ cat* $L = 2$. But a decomposition $\epsilon\, \mathfrak{H} - C^*(L)$ is clearly $\{L_1, L_2\}$ where $L_1 = I_1 + \sum_{u=1}^{2} D_u + T + \sum_{u=-2}^{2} W_u + \sum_{u=-2}^{-1} D_u$ and $L_2 = \sum_{u=-1}^{1} I_u$.

V. Miscellany

42. Generalizations. In the preceding chapters no attempt was made to define the most general category. We have considered a class of categories which is natural and has a meaningful theory. I shall now indicate several generalizations which are natural extensions.

First: Let K be a neighborhood retract of the arcwise connected space M; a set A is categorical rel K in M if there is an open set containing A which can be deformed in M into K; a covering σ of X belongs to $C_M(X \text{ rel } K)$ if each set of σ is categorical rel K in M. The minimum, $\text{cat}_M\,(X \text{ rel } K)$, of $|\sigma|$ as σ ranges over $C_M(X \text{ rel } K)$ has many properties similar to those of $\text{cat}_M X$ to which it reduces if K is contractible in M.

We note several properties of $\text{cat}_M(X \text{ rel } K)$:

If M is an absolute neighborhood retract, a closed set is categorical rel K if and only if it can be deformed in M into K.

If X is closed and M is an absolute neighborhood retract then $\text{cat}_M\,(X \text{ rel } K) \leqq 1 + \dim\,(X \cdot \overline{M - K})$.

By the method of theorem 20.1 it can be shown that if $M = M_1 \times \cdots \times M_k$ is an absolute neighborhood retract and T denotes the set of points $x = (x_1, \cdots, x_k)$ of M which have at least one coordinate identical with the corresponding coordinate of a fixed point $p = (p_1, \cdots, p_n)$, then $k\, \text{cat}_M\,(M \text{ rel } T) \leqq \text{cat}\, M + k - 1$.

This last implies theorem 20.1 because if $\text{cat}\, M_i \geqq 2$ for $i = 1, \cdots, k$ and M is essential then $\text{cat}_M\,(M \text{ rel } T) \geqq 2$.

Second: Let Φ be a subset of W^M. A covering σ of X by open sets of M belongs to $C_M(X \,\|\, \Phi)$ if $\phi \mid A$ is homotopic to a constant for every $\phi \,\epsilon\, \Phi$ and every $A \,\epsilon\, \sigma$. As usual $\text{cat}_M\,(X \,\|\, \Phi)$ is the minimum of $|\sigma|$ as σ ranges over $C_M(X \,\|\, \Phi)$. Clearly $\text{cat}_M\,(X \,\|\, \Phi) = \text{cat}_M X$ when $W = M$ and Φ consists of the single mapping 1.

In order that a family $\Phi \,\epsilon\, S_1^M$ be k-compatible [20, p. 158] it is necessary and sufficient that $\text{cat}_M\,(M \,\|\, \Phi) \leqq k$. Thus [20, p. 180] the multicoherence $r_k(M) \geqq n$ if and only if there exist n linearly independent mappings $f_1, f_2, \cdots, f_n \,\epsilon\, S_1^M$ such that $\text{cat}_M\,(M \,\|\, \{f_1, \cdots, f_n\}) \leqq k$.

Furthermore [20, p. 180, theorem 1, and p. 188, theorem 1] $H_1\, \text{cat}\,(M, \mathfrak{R}_1) = \text{cat}_M\,(M \,\|\, S_1^M)$, and [20, p. 188, theorem 8] $\text{cat}_M\,(M \,\|\, S_1^M) \leqq 1 +$ maximum number of linearly independent mappings of S_1^M.

Princeton, New Jersey.

BIBLIOGRAPHY

(1) Alexander, J. W., An example of a simply connected surface bounding a region which is not simply connected. Proc. Nat. Acad. 10 (1924), 8–10.

(2) Alexandroff, P., and Hopf, H., Topologie I. Springer (1935), Berlin.

(3) Aronzajn, N., and Borsuk, K., Sur la somme et le produit combinatoire des rétractes absolus. Fund. Math. 18 (1932), 193–197.

(4) Bassi, A., Su alcuni nuovi invarianti delle varietá topologische. Annali di Math. 16 (1935), 275–297. Abstracted: On some new invariants of a manifold. Proc. Nat. Acad. 22 (1936), 698–699.

(5) Birkhoff, Garrett, Lattices and their applications. Bull. of the American Math. Soc. 44 (1938), 793–800.

(6) Borsuk, K., Sur les rétractes. Fund. Math. 17 (1931), 152–170.

(7) Borsuk, K., Über eine Klasse von lokal zusammenhängenden Räumen. Fund. Math. 19 (1932), 220–242.

(8) Borsuk, K., Drei Sätze über die n-dimensionale euklidische Sphäre. Fund. Math. 20 (1933), 177–190.

(9) Borsuk, K., and Mazurkiewicz, S., Sur les rétractes absolus indecomposables. Comptes Rend. 199 (1934), 110–112.

(10) Borsuk, K., Sur la décomposition des courbes régulières en dendrites. Fund. Math. 22 (1934), 287–291.

(11) Borsuk, K., Quelques rétractes singuliers. Fund. Math. 24 (1935), 249–258.

(12) Borsuk, K., Contribution à la topologie des polytopes. Fund. Math. 25 (1935), 51–58.

(13) Borsuk, K., Über das Phänomen der Unzerlegbarkeit in der Polyedertopologie. Comm. Math. Helvetici 8 (1935), 142–148.

(14) Borsuk, K., Über den Lusternik-Schnirelmannschen Begriff der Kategorie. Fund. Math. 26 (1936), 123–136. Abstracted: Sur la notion de la catégorie de M. L. Lusternik et Schnirelmann. Comptes Rend. 198 (1934), 1731–1732.

(15) Borsuk, K., Sur la décomposition des polyèdres n-dimensionnels en polyèdres contractiles en soi. Comp. Math. 3 (1936), 431–434.

(16) Denjoy, A., and Wolff, J., Sur la division d'une sphère en trois ensembles. Ens. Math. 32 (1933), 66–68.

(17) Eilenberg, S., Sur quelques propriétés topologiques de la surface de sphère. Fund. Math. 25 (1935), 267–272.

(18) Eilenberg, S., Sur la théorème de décomposition de la theorie de la dimension. Fund. Math. 26 (1936), 146–149.

(19) Eilenberg, S., Sur un théorème topologique de M. L. Schnirelmann. Math. Sbornik N.S. 1 (1936), 557–559.

(20) Eilenberg, S., Sur les espaces multicohérents I. Fund. Math. 27 (1936), 153–190.

(21) Eilenberg, S., Un théorème sur l'homotopie. Ann. of Math. 38 (1937), 656–661.

(22) Eilenberg, S., Sur les transformations à petites tranches. Fund. Math. 30 (1938), 92–95.

(23) Hirsch, G., Une généralisation d'un théorème de M. Borsuk concernant certaines transformations de l'Analysis Situs. Bull. Acad. Roy. Belg., 23, (1937), 219–225.

(24) Hopf, H., Freie Überdeckungen und freie Abbildungen. Fund. Math. 28, (1937) 33–57.

(25) Hurewicz, W., Topologie der Deformationen I. Proc. Akad. Amsterdam 38 (1935), 112–119.

(26) Hurewicz, W., Topologie der Deformationen II. Proc. Akad. Amsterdam 38 (1935), 521–528.

(27) Hurewicz, W., Topologie der Deformationen III. Proc. Akad. Amsterdam 39 (1936), 117–126.

(28) Hurewicz, W., Topologie der Deformationen IV. Proc. Akad. Amsterdam 39 (1936), 215–224.

(29) van Kampen, E., On the connection between the fundamental groups of some related spaces. American Journal of Math. 55 (1933), 261-267.
(30) Kuratowski, C., Topologie I. Monografje Matematyczne III (1933), Warsawa-Lwów.
(31) Kuratowski, C., Sur les espaces localement connexes et péaniens en dimension n. Fund. Math. 24 (1935), 269-287.
(32) László, R., Zu einen geometrischen Satz von K. Borsuk. Math. Fiz. Lap. 41 (1934), 36-40.
(33) Lefschetz, S., Topology. American Math. Soc. Colloquium Publications (1930), New York.
(34) Lefschetz, S., On locally connected and related sets. Ann. of Math. 35 (1934), 118-129.
(35) Lefschetz, S., On locally connected sets and retracts. Proc. Nat. Acad. 24 (1938), 392-393.
(36) Lusternik, L., and Schnirelmann, L., Sur un principe topologique en analyse. Comptes Rend. 188 (1929), 295-297.
(37) Lusternik, L., and Schnirelmann, L., Existence de trois lignés géodésiques fermées sur toute surface de genre 0. Comptes Rend. 188 (1929), 534-536, and 189 (1929), 269-271.
(38) Lusternik, L., Topologische Grundlagen der allgemeinen Eigenwerttheorie. Monatshefte für Math. u. Physik 37 (1930), 125-130.
(39) Lusternik, L., and Schnirelmann, L., Méthodes topologiques dans les problèmes variationnels. Inst. for Math. and Mechanics (1930), Mosçau. (in russian).
(40) Lusternik, L., and Schnirelmann, L., Méthodes topologiques dans les problèmes variationnels. Actualités scientifiques et industrielles (1934), Paris.
(41) Pontrjagin, L., Topological Groups. Princeton U. Press (1939) Princeton (translated from the Russian).
(42) Schnirelmann, L., Über eine neue kombinatorische Invariante. Monatshefte für Math. u. Physik 37 (1930), 131-134.
(43) Seifert, H., and Threlfall, W., Lehrbuch der Topologie. Teubner (1934), Leipzig.
(44) Wilder, R. L., On the imbedding of subsets of a metric space in Jordan continua. Fund. Math. 19 (1932), 45-64.
(a) Froloff, S., and Elsholz, L., Limite inférieure pour le nombre des valeurs critiques d'une fonction, donnée sur une variété. Math. Sbornik 42 (1935), 637-643.
(b) Gordon, J., On the minimal number of critical points of a real function defined on a manifold. Math. Sbornik N.S. 4 (1938), 105-113.
(c) Elsholz, L., Zur Theorie der Invarianten, die zur Bestimmung der unteren Grenze der Anzahl der kritischen Punkte einer stetigen Funktion, die auf einer Mannigfaltigkeit bestimmt ist, dienen können. Math. Sbornik N.S. 5 (1939), 551-558 (in russian).
(d) Elsholz, L., Die Länge einer Mannigfaltigkeit und ihre Eigenschaften. Math. Sbornik N.S. 5 (1939), 565-571 (in russian).
(e) Fox, R. H., On Homotopy and Extension of Mappings. Proc. Nat. Acad. 26 (1940), 26-28.
(f) Seifert, H., and Threlfall, W. Variationsrechnung in Grossen. Teubner (1938) Leipzig.

Annals of Mathematics
Vol. 42, No. 2, April, 1941

ANALYTIC CURVES

By Joachim Weyl

(Received August 14, 1939)

Introduction

An algebraic curve in the complex k-dimensional space $\mathfrak{R}$: $\{x_0, x_1, \cdots, x_k\}$ can be parametrically represented by setting up $x_1/x_0, x_2/x_0, \cdots, x_k/x_0$ as functions of rational character on a closed Riemann surface, the parameter being a point on the latter. From this viewpoint the algebraic curve appears as the realization of an abstract Riemann surface.

Our intention is to investigate curves $\mathfrak{C}$, defined parametrically in $\mathfrak{R}$ by setting up $x_1/x_0, x_2/x_0, \cdots, x_k/x_0$ as functions of rational character on an arbitrary Riemann surface $\mathfrak{F}$.

The methods used to accomplish this aim are based essentially upon R. Nevanlinna's theory of meromorphic functions in the generalized form developed for the investigation of meromorphic curves: H. and J. Weyl; Ann. of Math. vol. 39, pg. 516 (1938). This paper, of which the present endeavour is but an extension, will be quoted so frequently that in future it shall be referred to as "M C." In the present context the meromorphic curves appear as the special or—as we shall say—classical case where $\mathfrak{F}$ is the finite z-plane, hence $x_i = x_i(z)$ are meromorphic functions of z.

In the initial chapter we develop the first and second main theorems for the general case. In addition it contains a section concerned with the behaviour of the order of a realization of $\mathfrak{F}$ under (Kronecker-) multiplication with some other such realization, and under its projection into a lower-dimensional space. A final section discusses two specific examples; one where $\mathfrak{F}$ is an n-sheeted unbounded covering surface of the finite z-plane (Algebroid Curves); the second one where $\mathfrak{F}$ is the doubly punctured z-sphere (Ring-meromorphic Curves).

The second chapter is devoted to the defect relations (Third Main Theorem), so called because they represent the generalization of that relation holding for meromorphic functions. Their validity is shown here, besides in the classical case, only for the two specific examples mentioned above. The real addition to our theory made in this part is a modification of these relations so as to make their validity independent of the hypothesis that the exceptional points satisfy no accidental linear relations, a restriction which previously had to be imposed.

The author wishes to express in this place his gratitude and deep indebtedness to Professor H. Weyl, whose benevolent advice and frequent encouragement were essential in the completion of this paper.

I. The First and Second Main Theorems

1. Preliminary Considerations

a.) *Concerning the space* $\mathfrak{R}$:

In connection with the k-dimensional projective space $\mathfrak{R}$ we shall make use of those concepts and relations between them which were developed for the purpose of investigating the meromorphic curves in $\mathfrak{R}$. The notation employed then will be used without changes in the present treatment. To avoid repetition I shall refer the reader to the part entitled: "Distances and Means in Projective Space" of the paper[1] mentioned previously, and to H. Weyl's note on unitary metrics in projective space[2].

b.) *Concerning the surface* $\mathfrak{F}$:

Let $\mathfrak{F}$ be an arbitrary Riemann surface. Let a compact part G_0 of the surface be designated as *conductor* and fixed once and for all.

The quantities to be defined in the course of this investigation will depend on a region G whose closure is compact and which contains G_0 in its interior. Such a G shall be referred to as an admissible region. If $\mathfrak{F}$ is compact then $\mathfrak{F}$ itself is an admissible region. Primary consideration will be given to the case where both G and G_0 are connected although the actual treatment allows application to other cases as well. It will be assumed that the boundaries Γ_0 and Γ of G_0 and G, respectively, consist of a finite number of simple closed curves whose tangents vary continuously.

We now think of $\mathfrak{F}$ as a *condenser* with G_0 as the inner, charged conductor and $\bar{G} = \mathfrak{F} - G$ as the outer, grounded one. The electrostatic potential $\Phi^*(\mathfrak{p})$, defined in every point $\mathfrak{p}$ of $\mathfrak{F}$, which arises if G_0 is kept at the potential 1 and $\bar{G}$ at the potential 0, is harmonic in the intermediate dielectric $G - G_0$ and continuous on the whole surface. It follows from the principle concerning the maximum and minimum of a harmonic function that

$$0 \leqq \Phi^*(\mathfrak{p}) \leqq 1$$

throughout $\mathfrak{F}$. Consequently the density of charge

$$\rho^* = -\frac{1}{2\pi}\frac{\partial\Phi^*}{\partial n},$$

where n designates the normal directed toward the interior of the condenser, is $\geqq 0$ on Γ_0 and $\leqq 0$ on Γ. Strictly speaking we describe ρ^* in the neighborhood of any point $\mathfrak{p}$ on Γ_0 or Γ by means of a local uniformizing-parameter t of $\mathfrak{F}$ at that point. The designation $\partial/\partial n$, and later on ds—the element of arc-length of Γ_0 or Γ at $\mathfrak{p}$—, have to be understood therefore as referring to the image in

[1] M C. pp. 516–520.

[2] H. Weyl: On Unitary Metrics in Projective Space, Ann. of Math., vol. 40, no. 1, p. 141.

the t-plane. The operation $(\partial/\partial n)\,ds$ however is independent of the choice of the local parameter. This applies in particular to

$$-\frac{1}{2\pi}\frac{\partial\Phi^*}{\partial n}\,ds = d\sigma^*,$$

which is the charge of an element ds of arc-length on Γ_0 or Γ. Changing the sign in the definition of $d\sigma^*$ for an element ds on Γ, we have

$$d\sigma^* \geqq 0 \qquad \text{on } \Gamma_0 \text{ and } \Gamma.$$

The quantity

$$\int_{\Gamma_0} d\sigma^* = e$$

is the charge of the inner conductor which creates the potential Φ^*. Since the outer conductor assumes inductively the same amount of negative charge we also have

$$\int_{\Gamma} d\sigma^* = e.$$

The total energy E used to build up the charge e in the conductor G_0 is equal to the Dirichlet integral $D(\Phi^*)$ of the potential Φ^* taken over $G - G_0$ (or the whole surface $\mathfrak{F}$):

$$E = D(\Phi^*) = -\frac{1}{2\pi}\int_{\Gamma_0} \Phi^*\frac{\partial\Phi^*}{\partial n}\,ds = e.$$

Since Φ^* is not constant over the whole surface, $D(\Phi^*)$ and therefore e are actually greater than zero. The constant C with which the potential difference across $G - G_0$ has to be multiplied to obtain the charge creating it is usually referred to as the *capacity* of the condenser $\mathfrak{F}$:

$$C(\Phi_{\Gamma_0} - \Phi_{\Gamma}) = e.$$

For the particular potential Φ^* we have $C = e > 0$. Therefore we can form

$$\Phi = C^{-1}\Phi^*,$$

which is the solution of the electrostatic problem for the condenser $\mathfrak{F}$ with a normalized charge equal to unity on the conductor G_0. We obtain

(1.1) $$\int_{\Gamma} d\sigma = \int_{\Gamma_0} d\sigma = 1$$

with

$$d\sigma = -\frac{1}{2\pi}\frac{\partial\Phi}{\partial n}\,ds \qquad \text{on } \Gamma_0,$$

$$d\sigma = +\frac{1}{2\pi}\frac{\partial\Phi}{\partial n}\,ds \qquad \text{on } \Gamma.$$

If the conductor G_0 consists of several connected parts we have to think of the latter as initially being connected with each other by thin wires which will be eliminated again after the total charge has reached a distribution in equilibrium.

Let $u(\mathfrak{p})$ be a harmonic function on $\mathfrak{F}$. We apply Green's formula

$$\oint \left(u \frac{\partial \Phi}{\partial n} - \Phi \frac{\partial u}{\partial n} \right) ds = 0 \tag{1.2}$$

to the whole surface $\mathfrak{F}$. Some caution however is required because $\partial\Phi/\partial n$ has a jump

$$\lim \left\{ \left(\frac{\partial \Phi}{\partial n} \right)_- - \left(\frac{\partial \Phi}{\partial n} \right)_+ \right\} = \left[\frac{\partial \Phi}{\partial n} \right]$$

across Γ_0 as well as across Γ. The indices $-$ and $+$ refer to the values of that function at two points lying—so to speak—on opposite banks of these curves. The functions u, $\partial u/\partial n$, and Φ itself remain continuous at these junctions. Application of (1.2) results in

$$\frac{1}{2\pi} \int_\Gamma \left[\frac{\partial \Phi}{\partial n} \right] u\, ds + \frac{1}{2\pi} \int_{\Gamma_0} \left[\frac{\partial \Phi}{\partial n} \right] u\, ds = 0$$

or, since

$$\left[\frac{\partial \Phi}{\partial n} \right] ds = 2\pi\, d\sigma \quad \text{on } \Gamma, \qquad \left[\frac{\partial \Phi}{\partial n} \right] ds = -2\pi\, d\sigma \quad \text{on } \Gamma_0,$$

in the relation

$$\int_\Gamma u\, d\sigma - \int_{\Gamma_0} u\, d\sigma = 0. \tag{1.3}$$

Next consider a function F, meromorphic on $\mathfrak{F}$, that is a function which in every admissible region is of rational character. Then $u = \log | F |$ is harmonic on $\mathfrak{F}$ except for isolated singularities which it has at those points where F either possesses a pole or a zero. Before applying (1.3) we therefore cut out these isolated points by means of small circles surrounding them whose radii we let converge toward zero later on. There are only a finite number of zeros and poles of F in any admissible region; for the latter form a closed set of isolated points. The resulting equation will be

$$\int_\Gamma \log | F |\, d\sigma - \int_{\Gamma_0} \log | F |\, d\sigma = \sum\nolimits_{\mathfrak{p}_0} \Phi(\mathfrak{p}_0) - \sum\nolimits_{\mathfrak{p}_\infty} \Phi(\mathfrak{p}_\infty), \tag{1.4}$$

in future referred to as the fundamental equation. The sums on the right will have to be extended over all zeros $\mathfrak{p}_0$, and poles $\mathfrak{p}_\infty$ respectively of F on $\mathfrak{F}$, each counted with its proper multiplicity. But only the finite number of them which are contained in G actually contribute anything to these sums.

c.) *Concerning the curve* $\mathfrak{C}$:

An analytic curve $\mathfrak{C}$ of class $\mathfrak{F}$ in $\mathfrak{R}$ shall be defined as follows:

DEFINITION: In each point $\mathfrak{p}$ of $\mathfrak{F}$ there are given $(k + 1)$ function elements

$$x_i(\mathfrak{p}) = c_i + c_i't + \cdots, \qquad [i = 0, 1, \cdots, k], \tag{1.5}$$

t being a local uniformizing parameter of $\mathfrak{F}$ in the point $\mathfrak{p}$, with the properties

1.) $(c_0, c_1, \cdots, c_k) \neq (0, 0, \cdots, 0)$.

2.) The elements in $\mathfrak{p}'$, a point sufficiently near $\mathfrak{p}$, are obtained from the $x_i(\mathfrak{p})$ by direct analytic continuation on $\mathfrak{F}$.

3.) Effecting upon the elements (1.5) the changes:

α.) Replacement of $x_i(\mathfrak{p})$ by $\rho(\mathfrak{p})x_i(\mathfrak{p})$, where the common factor $\rho(\mathfrak{p}) = \rho_0 + \rho_1 t + \cdots$ does not vanish at $\mathfrak{p}$: $t = 0$,

β.) Replacement of t by some other local uniformizing parameter τ

$$t = g_1\tau + g_2\tau^2 + \cdots, \; g_1 \neq 0,$$

does not alter the point $(c_0, c_1, \cdots, c_k)$ which they define in $\mathfrak{R}$.

This point shall be called the point $\mathfrak{p}$ of the curve $\mathfrak{C}$. Furthermore we assume that $\mathfrak{C}$ does not lie in any linear subspace of $\mathfrak{R}$.

This definition makes it appear natural to look at the surface $\mathfrak{F}$ as the curve in the abstract of which $\mathfrak{C}$ is a concrete realization in the k-dimensional projective space $\mathfrak{R}$.

In order to obtain a truly geometric description of $\mathfrak{C}$ we have to do this in terms of quantities and relations which are invariant not only regarding the changes α.) and β.) but also under

γ.) an arbitrary non-singular linear transformation of the projective coordinate system:

$$y_i = \textstyle\sum_j h_{ij}x_j, \qquad \det(h_{ij}) \neq 0.$$

A given linear form

$$(\alpha x) = \textstyle\sum_0^k \alpha_i x_i(\mathfrak{p}), \qquad (\alpha_0, \alpha_1, \cdots, \alpha_k) \neq (0, 0, \cdots, 0),$$

will vanish to a certain finite order $h = h(\mathfrak{p}; \alpha) \geqq 0$ at $\mathfrak{p}$ if developed into a power series of the local parameter t. We shall call h the multiplicity with which $\mathfrak{C}$ cuts the plane α: $(\alpha_0, \alpha_1, \cdots, \alpha_k)$ in the point $\mathfrak{p}$. It is not affected by α.) nor β.), and, for that matter, neither by γ.) because the plane coordinates α_i of α are transformed under the latter's influence contragrediently to the point coordinates x_i.

The determinants

$$[x(\mathfrak{p})x'(\mathfrak{p}) \cdots x^{(l-1)}(\mathfrak{p})]_{i_1 \cdots i_l} = \begin{vmatrix} x_{i_1}(\mathfrak{p}) & \cdots & x_{i_l}(\mathfrak{p}) \\ \vdots & & \vdots \\ x_{i_1}^{(l-1)}(\mathfrak{p}) & \cdots & x_{i_l}^{(l-1)}(\mathfrak{p}) \end{vmatrix} \tag{1.6}$$

where the prime denotes differentiation with respect to t, assume each the factor factor ρ^l under α.) while under β.) they take on the factor

$$(dt/d\tau)^{0+1+2+\cdots+(l-1)},$$

which again amounts to the multiplication with a gauge factor different from zero at $\mathfrak{p}:\tau = 0$. It follows that the elements

$$(1.7) \qquad x_{i_1\cdots i_l}(\mathfrak{p}) = c_{i_1\cdots i_l} + c'_{i_1\cdots i_l}t + \cdots, \qquad [i_1 < i_2 < \cdots < i_l],$$

obtained from the determinants (1.6) by removing all factors corresponding to possible common zeros:

$$[x(\mathfrak{p})x'(\mathfrak{p}) \cdots x^{(l-1)}(\mathfrak{p})]_{i_1\cdots i_l} = t^{d_l}x_{i_1\cdots i_l}(\mathfrak{p}),$$

possess properties 1.), 2.), and 3.); consequently they define a realization $\mathfrak{C}_l$ of $\mathfrak{F}$ in a space $\mathfrak{R}_l$ of $k_l = \binom{k+1}{l} - 1$ dimensions. $\mathfrak{C}_l$ is the curve $\mathfrak{C}$ defined as the locus of its generating $(l-1)$-spreads. The multiplicities $d_l = d_l(\mathfrak{p})$ are not altered by any of the changes α.), β.), and γ.). A detailed local investigation of the manifolds $\mathfrak{C}_l$ can now be carried through along the lines followed in M C, pp. 521–522.

2. The First Main Theorem

Let $(\alpha x) = 0$ and $(\beta x) = 0$ be two planes in $\mathfrak{R}$. Into these linear forms substitute the function $x_i(\mathfrak{p})$ defining a realization $\mathfrak{C}$ of $\mathfrak{F}$ in $\mathfrak{R}$. Then the function

$$F = (\alpha x(\mathfrak{p}))/(\beta x(\mathfrak{p}))$$

will be meromorphic on $\mathfrak{F}$. Hence we can apply to it the fundamental equation (1.4). We introduce the notation

$$N(G; \alpha) = \sum \Phi(\mathfrak{p}_0) = \sum h(\mathfrak{p}; \alpha)\Phi(\mathfrak{p}),$$

where the first sum is to be extended over all zeros $\mathfrak{p}_0$ of F, each counted with its proper multiplicity; the second one however over all points $\mathfrak{p}$ of the Riemann surface $\mathfrak{F}$. Furthermore we write

$$\int_\Gamma \log \| \alpha x \|^{-1} d\sigma = m(G; \alpha),$$

$$\int_{\Gamma_0} \log \| \alpha x \|^{-1} d\sigma = m^0(G; \alpha).$$

Then (1.4) states that

$$(2.1) \qquad T(G) = N(G; \alpha) + m(G; \alpha) - m^0(G; \alpha)$$

is independent of the particular plane α. This is the first main theorem for analytic curves[3]. It is to be remembered that

$$\Phi \geqq 0 \text{ on } \mathfrak{F}$$

$$d\sigma \geqq 0 \text{ on } \Gamma \text{ and } \Gamma_0 .$$

[3] A. Dinghas: Bemerkungen zur Ahlfors'schen Methode in der Theorie der meromorphen Funktionen, Comp. Math. 5. pp. 107–118, 1937. Dinghas generalized R. Nevanlinna's

From the latter it follows that the compensating terms $m(G; \alpha)$ and $m^0(G; \alpha)$ are $\geqq 0$. In the classical case of a meromorphic curve $m^0(G; \alpha)$ proved to be independent of $G: |z| < r$. In the general case this is no longer true, for the distribution of charge $d\sigma$ on Γ_0 will in general depend on G. m^0 may not even be bounded as a function of G. It will however be bounded when no intersections of $\mathfrak{C}$ with the plane α lie on Γ_0, and under these circumstances this is a consequence of its continuity as a function of $\alpha_0, \alpha_1, \cdots, \alpha_k$. In any case the normalization (1.1) of the charge on G_0 brings it about that the average over all planes α is

$$\mathfrak{M}_\alpha m^0(G; \alpha) = \mathfrak{M}_\alpha m(G; \alpha) = \text{const.}$$

independently of G.

The part $N(G; \alpha)$ is invariant under a transformation γ.) while the compensating terms take on under its influence an additive term lying between the bounds $\pm \log K$ independently of G, K^2 being the quotient of the maximum and minimum of the Hermitian form $\sum_0^k | \sum_0^k h_{ij} x_j |^2$ under the restriction $\sum_0^k |x_i|^2 = 1$. Let us call two functions $T_1(G)$ and $T_2(G)$ equivalent: $T_1(G) \sim T_2(G)$ if $| T_1(G) - T_2(G) |$ remains below a fixed bound for all admissible regions G. In the sense of this equivalence $T(G)$ is independent of the projective coordinate system. Thus it seems natural to say of two realizations of the curve $\mathfrak{F}$ (which are set in correspondence to each other by their co-parametrization through $\mathfrak{p}$) that they are of the same order if their T-functions are equivalent.

Next we apply the two averaging processes $\mathfrak{M}_\alpha$ and $\mathfrak{M}_\alpha^a$ [4] to the relation (2.1). Its left side, being independent of α, will not be affected at all. We shall formulate the resulting relations at once for all curves $\mathfrak{C}_l$ ($\mathfrak{C}_1 \equiv \mathfrak{C}$). Denoting by the subscript l that the quantities, thus marked, are to be formed in $\mathfrak{R}_l$ by means of the functions (1.7) as their prototypes were formed in $\mathfrak{R}$ by means of the functions (1.5), we introduce the conventions

$$\mathfrak{M}_\alpha N_l(G; \alpha) = N_l(G),$$

$$\mathfrak{M}_\alpha^a N_l(G; \alpha) = N_l^*(G; a),$$

$$\int_\Gamma \log [\![ax]\!]_l^{-1} \, d\sigma = m_l^*(G; a),$$

and note the relation

$$\mathfrak{M}_\alpha^a m_l(G; \alpha) - \mathfrak{M}_\alpha^a m_l^0(G; \alpha) = m_l^*(G; a) - m_l^{0*}(G; a).$$

---

characteristic $T(r)$ of a meromorphic function by weighting, for its computation, the points inside the circle $|z| < r$ with the values of a twice continuously differentiable, real-valued, but otherwise arbitrary function $\lambda(z)$, defined in every point of that circle. We preferred to generalize in the other direction, using reasonably arbitrary regions but weighting their points by the values of certain functions which are harmonic in their interior, thus avoiding the occurrence of the latter's Laplacian when applying Green's formula.

[4] M C. pp. 518–520.

Then the first main theorem appears in the form of the following three parallel expressions

$$
\begin{aligned}
T_l(G) &= N_l(G; \alpha) + m_l(G; \alpha) - m_l^0(G; \alpha), \\
T_l(G) &= N_l(G), \qquad\qquad (2.2)\\
T_l(G) &= N_l^*(G; a) + m_l^*(G; a) - m_l^{0*}(G; a).
\end{aligned}
$$

In what respects our argument so far, as well as the proofs of future results, will have to be modified in order to apply also to curves $\mathfrak{C}$ which are contained in a linear subspace of $\mathfrak{R}$ is a matter which has been discussed in the appendix of the previous paper[5]. This is of importance because the assumption that $\mathfrak{C}$ itself lies in no linear subspace of $\mathfrak{R}$ does not imply a similar behavior on the part of the curves $\mathfrak{C}_l$ with respect to the spaces $\mathfrak{R}_l$.

Let there be given two permissible regions G_1 and $G_2 : G_1 \subset G_2$, and let Φ_1^* and Φ_2^* be the corresponding potentials. Then $\Phi_1^* \leqq \Phi_2^*$ because the difference $\Phi_2^* - \Phi_1^*$ is

firstly $\geqq 0$ in the complimentary region $\bar{G}_1$,
secondly $= 0$ in G_0,

hence thirdly $\geqq 0$ on the boundary of the intermediate region $G_1 - G_0$, in the interior of which it is a regular harmonic function.

Therefore the same inequality holds throughout $G_1 - G_0$. Consequently we have in each point $\mathfrak{p}$

$$C_1\Phi_1(\mathfrak{p}) \leqq C_2\Phi_2(\mathfrak{p})$$

hence

$$C_1N(G_1 ; \alpha) \leqq C_2N(G_2 ; \alpha) \tag{2.3}$$

or

$$\begin{vmatrix} N(G_1; \alpha) & N(G_2; \alpha) \\ C_1^{-1} & C_2^{-1} \end{vmatrix} \leqq 0.$$

From $\Phi_1^* \leqq \Phi_2^*$ it follows moreover that Φ_1^* decreases more rapidly in passing from a point on Γ_0 (where both of them are equal to unity) towards the interior of the intermediate region. Thus

$$-\frac{\partial\Phi_1^*}{\partial n} \geqq -\frac{\partial\Phi_2^*}{\partial n} \qquad \text{and} \qquad d\sigma_1^* \geqq d\sigma_2^*$$

on Γ_0, therefore $e_1 \geqq e_2$ or $C_1 \geqq C_2$. This last inequality together with (2.3) yields

$$N(G_1 ; \alpha) \leqq N(G_2 ; \alpha)$$

whence follow the same inequalities for $T(G)$ by averaging over all planes α.

[5] M C. p. 537.

$$\begin{vmatrix} T(G_1) & T(G_2) \\ C_1^{-1} & C_2^{-1} \end{vmatrix} \leqq 0 \qquad \text{and a forteriori} \tag{2.4}$$

$$T(G_1) \leqq T(G_2).$$

In the classical case, where $\mathfrak{F}$ is the finite z-plane and G_0 and G are circles of radii r_0 and $r > r_0$ respectively around the origin, $T(G) \equiv T(r)$ was found to be an increasing convex function of log r. (A function of regular type.) The generalization of the first part of this description is contained in the second one of the inequalities (2.4). The convexity with respect to log r leads in the classical case to the more stringent inequality

$$\begin{vmatrix} T(G_1) & T(G_2) & T(G_3) \\ C_1^{-1} & C_2^{-1} & C_3^{-1} \\ 1 & 1 & 1 \end{vmatrix} \leqq 0$$

for three circular regions $G_1 \subset G_2 \subset G_3$. An investigation of its validity in the general case would meet with considerable difficulty.

Complete analogy between the classical and the general case will be obtained if we exhaust $\mathfrak{F}$ by means of a sequence of admissible regions G_r. This means that r is a real parameter and that the sequence $\{G_r\}$ has the following properties:

1.) $G_r \subset G_{r'}$ if $r' > r$.

2.) If $\mathfrak{p}$ be a point of $\mathfrak{F}$ then there exists an r such that $\mathfrak{p}$ is in $G_{r'}$ for all $r' > r$.

Now we define $T(r) \equiv T(G_r)$ as the order of the curve $\mathfrak{C}$. In general $T(r)$ depends on the way we exhaust $\mathfrak{F}$ by regions G_r. This applies in particular to the classical case where it might have appeared as though one studied the curve in a specific parametrization represented by the parameter z. But for the fundamental equation (1.4) the choice of parametrization proves of no importance since this relation is invariant under one-to-one conformal transformations of the underlying Riemann surface $\mathfrak{F}$.

If $\mathfrak{F}$ is closed, and therefore compact, the curve $\mathfrak{C}$ is algebraic. No exhaustion is necessary, the outer conductor can be dispensed with, Φ^* is identically equal to unity, and the first main theorem states that the number of intersections of a plane α with the curve $\mathfrak{C}$ is independent of α.

In (1.4) it is permissible to let the inner conductor G_0 shrink to a point. The resulting form of (1.4) was generally used in its stead by previous authors on this subject.

3. Products and Projections

The simple results derived in the following sections are new also for the classical case. To prove them at once for the case of analytic curves brings about no additional difficulties. They will strengthen us in the belief that the order $T(G)$ which we introduced is really a natural concept and has all the properties which we expect from an "order."

a.) *Products*:

The result obtained in this section is essentially a generalization of the fact that two algebraic curves of orders m and n respectively intersect in mn points.

Consider two realizations $\mathfrak{C}$ and $\mathfrak{D}$ of the same abstract curve $\mathfrak{F}$, one of which lies in a k-space, the other one in an h-space.

$$\mathfrak{C}: x_0:x_1:x_2: \cdots :x_k = x_0(\mathfrak{p}):x_1(\mathfrak{p}): \cdots :x_k(\mathfrak{p}),$$

$$\mathfrak{D}: y_0:y_1: \cdots :y_h = y_0(\mathfrak{p}):y_1(\mathfrak{p}): \cdots :y_h(\mathfrak{p}).$$

By Kronecker-multiplication we obtain a third realization

$$z_{ij}(\mathfrak{p}) = x_i(\mathfrak{p})y_j(\mathfrak{p})$$

in a space of $(kh + k + h)$ dimensions. We shall call it the direct product $\mathfrak{C} \times \mathfrak{D}$ of the realizations $\mathfrak{C}$ and $\mathfrak{D}$.

Now let us compare the orders of these three curves with each other. We shall change our notation and for the order $T(G)$ of $\mathfrak{C}$ we shall write $T[\mathfrak{C}]$ assuming that thus denoted orders of different realizations are taken with respect to the same admissible region.

For the computation of $T[\mathfrak{C}]$, $T[\mathfrak{D}]$ and $T[\mathfrak{C} \times \mathfrak{D}]$ we choose the planes

$$\alpha: (\alpha_0, \alpha_1, \cdots, \alpha_k) \qquad \text{in the } k\text{-space},$$

$$\beta: (\beta_0, \beta_1, \cdots, \beta_h) \qquad \text{in the } h\text{-space},$$

and

$$\gamma: \gamma_{ij} = \alpha_i\beta_j \qquad \text{in the product-space}$$

respectively. Calculating each order for the same region G we readily obtain from

$$\left(\sum_0^k \alpha_i x_i\right)\left(\sum_0^k \beta_j x_j\right) = \sum_{ij} \gamma_{ij} z_{ij},$$

$$\left(\sum |x_i|^2\right)\left(\sum |y_j|^2\right) = \sum_{ij} |z_{ij}|^2, \qquad \left(\sum |\alpha_i|^2\right)\left(\sum |\beta_i|^2\right) = \sum_{ij} |\gamma_{ij}|^2,$$

the relations

$$N[\mathfrak{C} \times \mathfrak{D}] = N[\mathfrak{C}] + N[\mathfrak{D}],$$

$$m[\mathfrak{C} \times \mathfrak{D}] = m[\mathfrak{C}] + m[\mathfrak{D}].$$

Hence

$$T[\mathfrak{C}] + T[\mathfrak{D}] = T[\mathfrak{C} \times \mathfrak{D}]. \tag{3.1}$$

The order of a direct product of realizations of an abstract curve $\mathfrak{F}$ is the sum of the orders of the factors.

A similar relation does not hold for the orders of higher rank.

b.) *Powers*:

We apply the theorem (3.1) to the n times reiterated product of the curve $\mathfrak{C}$ with itself. Let us define this realization $\mathfrak{C}^n$ of $\mathfrak{F}$ by the functions

$$x_{m_0m_1\cdots m_k}(\mathfrak{p}) = \left\{\frac{n!}{m_0!\,m_1!\,\cdots\,m_k!}\right\}^{\frac{1}{2}} x_0^{m_0}(\mathfrak{p}) \cdots x_k^{m_k}(\mathfrak{p}), \tag{3.2}$$

where $m_0, m_1, \cdots, m_k$ run over all non-negative integers whose sum $\sum_0^k m_i$ equals n. (The advantage of this choice over $x_{m_0\cdots m_k}(\mathfrak{p}) = x_0^{m_0}(\mathfrak{p}) \cdots x_k^{m_k}(\mathfrak{p})$ will become apparent as the discussion proceeds.) We calculate the order of $\mathfrak{C}$ for a plane $(\alpha x) = 0$. On the other hand we remark that $[(\alpha x)]^n = [\sum_0^k \alpha_i x_i]^n$ is a linear combination of the monomials (3.2) with the coefficients

$$\alpha_{m_0m_1\cdots m_k} = \left\{\frac{n!}{m_0!\,m_1!\,\cdots\,m_k!}\right\}^{\frac{1}{2}} \alpha_0^{m_0}\alpha_1^{m_1}\cdots\alpha_k^{m_k}.$$

Making use of this for the computation of $T[\mathfrak{C}^n]$ we obtain at once

$$N[\mathfrak{C}^n] = nN[\mathfrak{C}]$$

and

$$m[\mathfrak{C}^n] = nm[\mathfrak{C}]$$

since

$$\sum |\alpha_{m_0m_1\cdots m_k}|^2 = \{\textstyle\sum_0^k |\alpha_i|^2\}^n,$$

$$\sum |x_{m_0m_1\cdots m_k}|^2 = \{\textstyle\sum_0^k |x_i|^2\}^n.$$

The result

$$T[\mathfrak{C}^n] = nT[\mathfrak{C}]$$

permits the following interpretation: Every plane

$$\sum \alpha_{m_0m_1\cdots m_k} x_{m_0m_1\cdots m_k} = 0$$

in the product-space can at the same time be thought of as an algebraic surface of order n in the space $\mathfrak{R}$ defined by

$$\sum \alpha_{m_0m_1\cdots m_k} x_0^{m_0} x_1^{m_1} \cdots x_k^{m_k} = 0.$$

In either case the sum is to be extended over all sequences of non-negative integers $m_0, m_1, \cdots, m_k$ whose sum is n. Viewed in this light $T[\mathfrak{C}^n]$ is seen to describe the behaviour of $\mathfrak{C}$ with respect to an algebraic surface of order n in the same sense that $T[\mathfrak{C}]$ describes its behaviour with respect to a plane. We found that: *The order of $\mathfrak{C}$ referred to an algebraic suface of order n is equal to n times the order of $\mathfrak{C}$ (referred to a plane).* This implies for instance that the average density of intersections of $\mathfrak{C}$ with an algebraic surface of order n is n times the average density of its intersections with a plane. (Without the numerical factor $\{n!/m_0!m_1! \cdots m_k!\}^{\frac{1}{2}}$ the final result would have appeared less sharply in the form of an equivalence $T[\mathfrak{C}^n] \sim nT[\mathfrak{C}]$.)

c.) *Projections*:

The projection from a given $(k - n - 1)$-dimensional linear subspace $\mathfrak{R}^*$ of $\mathfrak{R}$ as center into a linear subspace $\mathfrak{R}'$ of n dimensions which does not intersect $\mathfrak{R}^*$ can be described in a suitably chosen projective coordinate system by the

passage from the point $(x_0, x_1, \cdots, x_n, x_{n+1}, \cdots, x_k)$ to the point $(x_0, x_1, \cdots, x_n, 0, \cdots, 0)$.

Under the influence of such a projection the curve $\mathfrak{C}$ passes into a curve $\mathfrak{C}'$ in $\mathfrak{R}'$. If a point $\mathfrak{p}$ of the original curve lies on $\mathfrak{R}^*$ then the elements

$$x_0(t), x_1(t), \cdots, x_n(t) \tag{3.3}$$

will all contain a common factor $t^b(b = b(\mathfrak{p}))$; only after its elimination do the power series (3.3) define the corresponding element of the projected curve. We consider a plane α of the special form

$$\alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Then it follows that

$$h'(\mathfrak{p}; \alpha) = h(\mathfrak{p}; \alpha) - b(\mathfrak{p}),$$

where h refers to $\mathfrak{C}$ and h' to $\mathfrak{C}'$. Consequently

$$h'(\mathfrak{p}; \alpha) \leqq h(\mathfrak{p}; \alpha)$$

and also

$$N'(G; \alpha) \leqq N(G; \alpha).$$

The same inequality for the m-part follows directly from the fact that

$$\frac{\sum_0^n |\alpha_j|^2 \sum_0^n |x_j|^2}{|\sum_0^n \alpha_j x_j|^2} = \left(\frac{1}{\| \alpha x \|'}\right)^2 \leqq \left(\frac{1}{\| \alpha x \|}\right)^2 = \frac{\sum_0^n |\alpha_j|^2 \sum_0^k |x_i|^2}{|\sum_0^n \alpha_j x_j|^2}$$

yielding

$$T' \leqq T + (m^0 - m^{0\prime}).$$

With $X^2 = \sum_0^k |x_i|^2$, $X'^2 = \sum_0^n |x_j|^2$, we have

$$m^0 - m^{0\prime} = \int_{\Gamma_0} \log (X : X')\, d\sigma$$

and the final result now appears in the form

$$T\{\mathfrak{R}'\} \leqq T\{\mathfrak{R}\} + \int_{\Gamma_0} \log (X : X')\, d\sigma \tag{3.4}$$

which is independent of the particular choice of the plane α. The notation used explains itself.

The additional term on the right side is certainly bounded as a function of G if none of the points which $\mathfrak{C}$ has in common with $\mathfrak{R}^*$ lie on Γ_0. In the classical case the independence of m^0 from G prevents this difficulty from arising altogether. The contents of the inequality (3.4) can be expressed with due qualifications in the form: *Projection cannot essentially increase the order of an analytic curve.*

The relation (3.4) has its counter-part for higher l. For our projection $\mathfrak{C}'$ we can form the l-curves $(\mathfrak{C}')_l$. They will lie in spaces $\mathfrak{R}'_l$ of dimensionality

$\binom{n+1}{l} - 1$, $(l \leqq n)$. We may think of these spaces $\mathfrak{R}'_l$ as linear subspaces of the corresponding $\mathfrak{R}_l$ respectively. The projection of $\mathfrak{R}$ from $\mathfrak{R}^*$ induces for higher l the projection of $\mathfrak{R}_l$ from $\mathfrak{R}_l^*$ which is effected through replacing in the determinants

$$[xx' \cdots x^{(l-1)}]_{i_1 \cdots i_l}$$

all $(k + 1)$-uples $x, x', \cdots$ by the corresponding ones with zeros after the $(n + 1)$st place. This projection throws $\mathfrak{C}_l$ into $\mathfrak{R}'_l$ where it is seen to coincide with $(\mathfrak{C}')_l$: The l-curves $(l \leqq n)$ of the projection are projections of the corresponding $\mathfrak{C}_l$.

Hence writing

$$X_l^2 = \sum | x_{i_1 \cdots i_l} |^2 \qquad [i_1 < i_2 < \cdots < i_l]$$

and putting $X_l'^2$ equal to what arises from X_l^2 under the projection of $\mathfrak{R}_l$ from $\mathfrak{R}_l^*$, induced by the production of $\mathfrak{R}$ from $\mathfrak{R}^*$, we obtain also

$$T\{\mathfrak{R}'_l\} \leqq T\{\mathfrak{R}_l\} + \int_{\Gamma_0} \log (X_l : X'_l) \, d\sigma$$

which by its very form proves to be independent not only of any particular plane but also—in the sense of equivalence—of the particular coordinate system employed to derive it.

We are led to a kind of inverse of the above relations when we consider the following situation: By projection from the centers $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$ the total space $\mathfrak{R}$ be mapped into $\mathfrak{R}'$ and $\mathfrak{R}''$ respectively. By $\mathfrak{R}' \vee \mathfrak{R}''$ and $\mathfrak{R}' \wedge \mathfrak{R}''$ let us denote the spaces into which $\mathfrak{R}$ turns by projection from the centers $\mathfrak{R}^* \cap \mathfrak{R}^{**}$ (intersection of $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$) and $\mathfrak{R}^* \cup \mathfrak{R}^{**}$ (sum, union of $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$) respectively. We assume in particular that $\mathfrak{R}^* \cap \mathfrak{R}^{**}$ is empty, then

$$\mathfrak{R}' \vee \mathfrak{R}'' = \mathfrak{R} \qquad \text{and we have}$$

$$\dim (\mathfrak{R}' \wedge \mathfrak{R}'') + \dim (\mathfrak{R}' \vee \mathfrak{R}'') = \dim (\mathfrak{R}') + \dim (\mathfrak{R}'').$$

We shall show that the relation

$$T\{\mathfrak{R}' \wedge \mathfrak{R}''\} + T\{\mathfrak{R}' \vee \mathfrak{R}''\} \leqq T\{\mathfrak{R}'\} + T\{\mathfrak{R}''\},$$

holding with proper qualifications, connects the corresponding orders. Two projections from the centers $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$ are called complimentary if not only

$$\mathfrak{R}^* \cap \mathfrak{R}^{**} = 0$$

but also

$$\dim (\mathfrak{R}^* \cup \mathfrak{R}^{**}) = \dim (\mathfrak{R}) - 1.$$

Then we have, since $\mathfrak{R}' \vee \mathfrak{R}'' = \mathfrak{R}$, the relation

$$\dim (\mathfrak{R}) = \dim (\mathfrak{R}') + \dim (\mathfrak{R}'')$$

and

$$T\{\mathfrak{R}\} \leqq T\{\mathfrak{R}'\} + T\{\mathfrak{R}''\}$$

The order of an analytic curve does not essentially exceed the sum of orders obtained when adding the orders of any two complimentary ones of its projections.

To prove this choose the coordinates so that the vanishing of

$$x_0, x_1, \cdots, x_r \tag{3.5}$$

defines $\mathfrak{R}^* \cup \mathfrak{R}^{**}$, whereas

$$x_0, x_1, \cdots, x_r, \qquad y_{r+1}, \cdots, y_s, \tag{3.6}$$

$$x_0, x_1, \cdots, x_r, \qquad z_{r+1}, \cdots, z_t, \tag{3.7}$$

are the coordinates associated in the same manner with $\mathfrak{R}^*$ and $\mathfrak{R}^{**}$ respectively. The sequence of the x, y, and z together is under these circumstances a full coordinate system of $\mathfrak{R}$. The quantities (3.5), (3.6), and (3.7) are the coordinates in $\mathfrak{R}^0 = \mathfrak{R}' \wedge \mathfrak{R}''$, $\mathfrak{R}'$, and $\mathfrak{R}''$ respectively. We choose a plane

$$\alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_r x_r = 0.$$

Then in the same manner as before

$$h^0 = h(\mathfrak{p}; \alpha) - b^0(\mathfrak{p}), \qquad h' = h(\mathfrak{p}; \alpha) - b'(\mathfrak{p}), \qquad h'' = h(\mathfrak{p}; \alpha) - b''(\mathfrak{p}).$$

Obviously both b' and b'' are $\leqq b^0$, moreover, since there is no point where all coordinates (3.6) and (3.7) are zero, one of the two numbers b' or b'' is necessarily zero, e.g. $b'' = 0$. In combining this with $b' \leqq b^0$ one gets the relation $b' + b'' \leqq b^0$ holding for both alternatives, consequently

$$h' + h'' \geqq h + h^0.$$

This leads to

$$N' + N'' \geqq N + N^0.$$

The corresponding relation for the m-part

$$m' + m'' \geqq m + m^0$$

follow at once from the inequality

$$X'^2 + X''^2 \geqq X^2 + X^{02},$$

where the quantities X^0, X', and X'' are the projections of X from the spaces $\mathfrak{R}^* \cup \mathfrak{R}^{**}$, $\mathfrak{R}^*$, and $\mathfrak{R}^{**}$ respectively. The exact form of the desired result is therefore given in the formula

$$T\{\mathfrak{R}'\} + T\{\mathfrak{R}''\} \geqq T\{\mathfrak{R}\} + T\{\mathfrak{R}' \wedge \mathfrak{R}''\} + \gamma^0$$

with

$$\gamma^0 = \int_{\Gamma_0} \log (X'X'' : XX^0)\, d\sigma,$$

which is independent of the particular plane used to derive it.

4. The Second Main Theorem

We first derive the second main theorem under the hypothesis that in addition to our curve $\mathfrak{C}: x_i = x_i(\mathfrak{p})$ we have a meromorphic function $z(\mathfrak{p})$ on $\mathfrak{F}$ at our disposal. Then $\mathfrak{F}$ may be looked upon as a covering surface of the z-plane. Another more appropriate interpretation would be to think of $z(\mathfrak{p})$ as a curve given in addition to $\mathfrak{C}$,—a realization of $\mathfrak{F}$ in a one-dimensional space. With the abbreviations

$$\alpha(xx' \cdots x^{(l-1)}) = \sum \alpha_{i_1 \cdots i_l}[xx' \cdots x^{(l-1)}]_{i_1 \cdots i_l}, \qquad [i_1 < i_2 < \cdots < i_l],$$

we form the expressions

$$F = \frac{\alpha(xx' \cdots x^{(l-2)}) \cdot \gamma(xx' \cdots x^{(l)})}{\{\beta(xx' \cdots x^{(l-1)})\}^2} \tag{4.1}$$

interpreting the prime as signifying derivation with respect to z. This is a meromorphic function on $\mathfrak{F}$. If we hold on to the convention that the prime indicates differentiation with respect to the local parameter t, then

$$F = \frac{\alpha(xx' \cdots x^{(l-2)}) \cdot \gamma(xx' \cdots x^{(l)})}{\{\beta(xx' \cdots x^{(l-1)})\}^2} \cdot \frac{dt}{dz}.$$

The expansion of dz/dt in terms of t at the point $\mathfrak{p}$ will start with a certain power t^j $(j = j(\mathfrak{p}))$. We introduce the quantity

$$\sum j(\mathfrak{p})\Phi(\mathfrak{p}) = J^z(G)$$

and the customary

$$V_l(G) = \sum \{d_{l+1}(\mathfrak{p}) - 2d_l(\mathfrak{p}) + d_{l-1}(\mathfrak{p})\}\Phi(\mathfrak{p}),$$

the sums extending over all points $\mathfrak{p}$ of $\mathfrak{F}$. The latter may be looked upon as a measure of the density of stationary $(l - 1)$-spreads of the realization $\mathfrak{C}$ over the part G of the abstract curve. Application of the fundamental equation (1.4) to the function yields in the same manner as in the classical case the second main theorem:

$$V_l(G) + \{T_{l+1}(G) - 2T_l(G) + T_{l-1}(G)\} = J^z(G) + \{{}^z\Omega_l(G) - {}^z\Omega_l^0(G)\} \tag{4.2}$$

where

$${}^z\Omega_l(G) = \int_\Gamma \log\left\{\left(\frac{X_{l+1}X_l}{X_l^2}\right)\left|\frac{dt}{dz}\right|\right\} d\sigma$$

and where ${}^z\Omega_l^0$ denotes the same integral extended over Γ_0. The expression under the integral sign is independent of both the gauge factor and the choice of the local parameter. The entire left side of (4.2) does not depend on the particular function z. The fact that the right member also has the same value for two meromophic function z and ζ is realized by applying the fundamental equation (1.4) to the function $d\zeta/dz$.

In order to obtain a formulation of the second main theorem which does not involve the auxiliary realization $z(\mathfrak{p})$ of $\mathfrak{F}$ we argue as follows: According to the theory of uniformization the surface $\mathfrak{F}$ is one of the spatial forms which the Euclidean, the spherical, or the Lobatschewskian plane may assume and is therefore endowed with a uniquely determined Riemannian line element $d\mathfrak{s}$. By means of this line element we can form the following expressions

$$\Omega_l(G) = \int_\Gamma \log\left(\frac{X_{l+1}X_{l-1}}{X_l^2}\cdot\frac{|dt|}{d\mathfrak{s}}\right)d\sigma$$

which depend neither on the gauge factor nor on the local parameter. Now consider two realizations $\mathfrak{C}$ and $\mathfrak{D}$ of $\mathfrak{F}$ and form the quotient of two expressions (4.1) for $\mathfrak{C}$ and $\mathfrak{D}$. We are not even forced to choose the same rank l in both cases. By application of the fundamental equation (1.4) to this quotient which does not depend on the local parameter we find that

$$V_l(G) + \{T_{l+1}(G) - 2T_l(G) + T_{l-1}(G)\} - \{\Omega_l(G) - \Omega_l^0(G)\} = \epsilon(G)$$

is independent of the choice of the curve $\mathfrak{C}$ as well as of l and therefore determined only by the Riemann surface $\mathfrak{F}$. From a theoretical standpoint this relation

$$V_l(G) + \{T_{l+1}(G) - 2T_l(G) + T_{l-1}(G)\} = \Omega_l(G) - \Omega_l^0(G) + \epsilon(G) \tag{4.3}$$

is the more satisfactory form of the second main theorem, although, for the purpose of our future estimates the equations (4.2) will prove the more useful of the two. If we apply (4.3) to $\mathfrak{C}$ and to the one-dimensional curve defined by z, we are led back by subtraction to the previous form.

In the algebraic case of a compact surface $\mathfrak{F}$ we shall take $G \equiv \mathfrak{F}$. Then the integrals along the boundaries vanish and we obtain the well-known Plücker formulae stating that

$$v_l + (n_{l+1} - 2n_l + n_{l-1})$$

is a constant independent of l and the curve.

5. Examples

A general discussion proceeding along the lines of the previous sections, which now would lead naturally to an estimate of first $\Omega_l(G)$ and then $\Omega_l(G)$ augmented by defect sums, meets with forbidding difficulties. We shall limit ourselves therefore to the discussion of two examples which possess the same rotational symmetry as the classical case and therefore allow description by polar coordinates. The regions G which are used to exhaust $\mathfrak{F}$ will be bounded by concentric circles. All quantities involved in our relations will become real-valued functions of the radii of those circles. This enables us to apply the methods of estimation which were employed in the classical case to derive the defect relations. For the moment let us specialize the results gained so far to fit the cases of the proposed examples.

a.) *Algebroid Curves*:

In this case $\mathfrak{F}$ is an n-sheeted covering surface over the open z-plane without relative boundaries. The conductor G_0 shall be that part of $\mathfrak{F}$ which covers the circle $|z| \leqq r_0$ of the z-plane. The parts G_r of $\mathfrak{F}$ which cover the circles $|z| < r$ make up the sequence $\{G_r\}$, used to exhaust $\mathfrak{F}$. A complete treatment of this case has been given, though from a different viewpoint, by E. Ullrich.[6]

In our case

$$\Phi(\mathfrak{p}) = \frac{1}{n} \log \frac{r}{r_0} \quad \text{for } \mathfrak{p} \text{ in } G_0,$$

$$\Phi(\mathfrak{p}) = \frac{1}{n} \log \frac{r}{|z|} \quad \text{for } \mathfrak{p} \text{ in } G_r \text{ (} z \text{ being its trace in the } z\text{-plane)},$$

$$\Phi(\mathfrak{p}) = 0 \quad \text{for } \mathfrak{p} \text{ in } \bar{G}_r,$$

and

$$d\sigma = \frac{1}{2\pi n} d\varphi.$$

Indicating by $\int_0^{2\pi n}$ the integration around the boundary of G, that is, writing

$$m_l(G; \alpha) = \frac{1}{2\pi n} \int_0^{2\pi n} \log \| \alpha x \|_l^{-1} \, d\varphi = m_l(r; \alpha)$$

and

$$m_l^*(G; a) = \frac{1}{2\pi n} \int_0^{2\pi n} \log [\![ax]\!]_l^{-1} \, d\varphi = m_l^*(r; a),$$

we obtain the first main theorem for algebroid curves in the form

$$T_l(r) \sim N_l(r; \alpha) + m_l(r; \alpha) \sim N_l(r) \sim N_l^*(r; a) + m_l^*(r; a)$$

because

$$m^0(r; \alpha) = m(r_0; \alpha),$$

formed by integrating around Γ_0, is independent of r and therefore bounded, which permits us to write the customary equivalences.

The second main theorem is best stated in the form (4.2) with z as the variable. We note immediately that ${}^z\Omega_l^0(G)$ is bounded for all G_r. Hence (4.2) appears, written as an equivalence, in the form

$$V_l(r) + \{T_{l+1}(r) - 2T_l(r) + T_{l-1}(r)\} \sim J(r) + \Omega_l(r).$$

[6] E. Ullrich: Wertverteilung und Verzweigtheit von Algebroiden, Crelle's Journal für die reine und angew. Mathematik, Bd. 167, p. 198.

For the quantity $\Omega_l(r) = {}^z\Omega_l(G_r)$, we have in keeping with the previously adopted notation,

$$\Omega_l(r) = \frac{1}{2\pi n}\int_0^{2\pi n} \log\left(\frac{X_{l+1}X_{l-1}}{X_l^2}\right) d\varphi,$$

while $J(r)$ measures the density of branchpoints of $\mathfrak{F}$ over the disc $|z| < r$ of the z-plane.

b.) *Ring-meromorphic Curves*:

As a second example we shall consider the curves $\mathfrak{C}$ defined by $(k+1)$ functions

$$x_0 = x_0(z), \qquad x_1 = x_1(z), \cdots, \qquad x_k = x_k(z),$$

which are meromorphic on the doubly punctured z-plane. Let us assume that $\mathfrak{F}$ is the z-plane punctured at $z = 0$ and $z = \infty$. The natural choice of our region G is a ring bounded by two concentric circles of radii R and r, $R > r$, respectively. The first one excludes a neighbourhood of $z = \infty$, the second one a neighbourhood of $z = 0$. Inside the region G we have to fix the conductor G_0 which is to contain the charge creating the potential Φ. We take, as the most convenient one, another ring bounded by the concentric circles of radii R_0 and r_0 respectively.

$$r < r_0 < R_0 < R.$$

So far the ring-meromorphic case escaped the careful treatment which has been given to the classical meromorphic as well as to the algebroid cases. The customary procedure of letting Φ be set up by a point charge destroys the rotational symmetry in the present case, thus complicating the situation considerably.

The region $G - G_0$ consists of two parts separated by the conductor G_0; one is bounded by the circles of radii R and R_0, let us call it ${}^eG$; the other one, ${}^iG$, by the circles of radii r and r_0. Their capacities are

$${}^eC = \left(\log\frac{R}{R_0}\right)^{-1}, \qquad {}^iC = \left(\log\frac{r_0}{r}\right)^{-1}.$$

Using these designations it is found that the capacity of the condenser $\mathfrak{F}$ is

$$C = {}^eC + {}^iC. \qquad \text{(Capacities connected in parallel.)}$$

Therefore we find for $\Phi(z)$ the following expressions

$$\Phi(z) = \{{}^eC/C\}\log\frac{R}{|z|} \qquad \text{for } z \text{ in } {}^eG,$$

$$\Phi(z) = \{{}^iC/C\}\log\frac{|z|}{r} \qquad \text{for } z \text{ in } {}^iG,$$

$$\Phi(z) = \{1/C\} \qquad \text{for } z \text{ in } G_0,$$

$$\Phi(z) = 0 \qquad \text{everywhere else.}$$

The charges carried by elements of arc-length on the various boundaries are

$$d\sigma = \frac{1}{2\pi}\,\{{}^eC/C\}\,d\varphi \quad \text{on the boundary of } {}^eG,$$

$$d\sigma = \frac{1}{2\pi}\,\{{}^iC/C\}\,d\varphi \quad \text{on the boundary of } {}^iG.$$

Writing down the relations (2.2) and (4.2) with these specializations it will be seen that all quantities involved show a characteristic pattern: They are the weighted averages of two quantities each having a similar significance, regarding one of the regions ${}^eG$ or ${}^iG$ alone, as the total terms have with regard to the whole region G. The weights are the capacities of the corresponding regions repectively. Thus

$$f(G) = [{}^ef({}^eG);\,{}^if({}^iG)] = \frac{{}^eC\cdot{}^ef + {}^iC\cdot{}^if}{C}. \tag{5.1}$$

For the compensating terms for example we have

$${}^em({}^eG;\alpha) = \frac{1}{2\pi}\int_0^{2\pi} \log \|\,\alpha x(Re^{i\varphi})\,\|^{-1}\,d\varphi = m(R;\alpha),$$

$${}^im({}^iG;\alpha) = \frac{1}{2\pi}\int_0^{2\pi} \log \|\,\alpha x(re^{i\varphi})\,\|^{-1}\,d\varphi = m(r;\alpha),$$

and similarly for m^0 and Ω^0. For the terms counting multiplicities let N serve as an example:

$${}^eN({}^eG;\alpha) = \sum_{R>|z|>R_0} h(z;\alpha)\log\frac{R}{|z|} + c\log\frac{R}{R_0} = {}^eN(R;\alpha)$$

$${}^iN({}^iG;\alpha) = \sum_{r<|z|<r_0} h(z;\alpha)\log\frac{|z|}{r} + c'\log\frac{r_0}{r} = {}^iN(r;\alpha)$$

where the constants c and c' must be chosen subject only to the condition

$$c + c' = \sum_{r_0\le|z|\le R_0} h(z;\alpha).$$

In other words: It is immaterial how the contributions to N, coming from intersections of $\mathfrak{C}$ with α in points over the conductor G_0, are divided up among ${}^eN$ and ${}^iN$. The resulting arbitrariness in the functions

$$\begin{aligned} {}^eT_l(r) &= {}^eN_l(r;\alpha) + m_l(r;\alpha) - m_l^0(r;\alpha) && (r > R_0) \\ {}^iT_l(r) &= {}^iN_l(r;\alpha) + m_l(r;\alpha) - m_l^0(r;\alpha) && (r < r_0) \end{aligned} \tag{5.2}$$

coincides with the one which the very structure of any expression (5.1) allows its constituents ${}^ef$ and ${}^if$. These components are not uniquely determined by f, for any pair

$${}^ef + c/{}^eC; \quad {}^if - c/{}^iC \qquad (c = \text{const.})$$

will perform equally well in their stead. The condition $f(G) > 0$ is satisfied by all quantities entering into the formulae of the first and second main theorems. We remark that under these circumstances a constant c can be found such that also

$$^{e}f^{*} = {}^{e}f + c/{}^{e}C > 0, \qquad {}^{i}f^{*} = {}^{i}f - c/{}^{i}C > 0.$$

Disregarding the positive denominator we are given two functions $f(x)$ and $g(y)$ such that for all values of x and y for which they are explained we have $f(x) + g(y) > 0$. We then have to find a constant c such that also

$$f^{*}(x) = f(x) + c > 0; \qquad g^{*}(y) = g(y) - c > 0.$$

Denoting by f and g the greatest lower bounds of $f(x)$ and $g(y)$ respectively, we see that $c = \frac{1}{2}(g - f)$ certainly has the required property, for

$$f^{*}(x) \geqq \tfrac{1}{2}(f + g) \geqq 0, \qquad g^{*}(y) \geqq \tfrac{1}{2}(f + g) \geqq 0,$$

where never more than one of the equality signs can possibly hold in any given case.

The specializations $r = r_0$ and $R = R_0$ respectively show that the first two main theorems for ring-meromorphic curves, being originally relations between weighted averages, permit the following formulation for the quantities in their unaveraged state: The functions (5.2), defined for sufficiently great and sufficiently small values of their arguments respectively, are independent of the particular plane α. They define in pairs the orders

$$T_l(G) = [{}^{e}T_l(R); {}^{i}T_l(r)]$$

of a ring-meromorphic curve in the sense that two such pairs are equivalent:

$$[{}^{e}T(R); {}^{i}T(r)] \sim [{}^{e}T^{*}(R); {}^{i}T^{*}(r)]$$

if

$$\begin{aligned} {}^{e}T^{*}(R) &= {}^{e}T(R) + c \log R + O(1), \\ {}^{i}T^{*}(r) &= {}^{i}T(r) + c \log r + O(1). \end{aligned}$$

With this same convention the second main theorem appears in the form

$$(5.3) \qquad \begin{aligned} {}^{e}V_l(r) + \{{}^{e}T_{l+1}(r) - {}^{e}T_l(r) + {}^{e}T_{l-1}(r)\} &= \Omega_l(r) - \Omega_l(R_0) \quad \text{for } r > R_0, \\ {}^{i}V_l(r) + \{{}^{i}T_{l+1}(r) - {}^{i}T_l(r) + {}^{i}T_{l-1}(r)\} &= \Omega_l(r) - \Omega_l(r_0) \quad \text{for } r < r_0. \end{aligned}$$

We finally remark that, since

$${}^{e}m^{0}({}^{e}G; \alpha) = m(R_0 ; \alpha) \sim 0; \qquad {}^{i}m^{0} = m(r_0 ; \alpha) \sim 0$$

we have also

$$m^{0}(G; \alpha) = [m(R_0 ; \alpha); m(r_0 ; \alpha)] \sim 0$$

and similarly

$$\Omega_l^{0}(G) = [\Omega_l(R_0); \Omega_l(r_0)] \sim 0$$

allowing us to replace the equality signs in (5.2) and (5.3) respectively by equivalences and dropping the terms m_l^0 and Ω_l^0.

II. The Third Main Theorem

6. General Relations

Some of the formulae, leading up to the estimates which finally culminate in the defect relations, can be derived in the general case. To do this is the aim of this first section.

Again we consider the function

$$w(\mathfrak{p}) = (\alpha x(\mathfrak{p}))/(\beta x(\mathfrak{p})) = w = w_1/w_2$$

which is meromorphic on $\mathfrak{F}$. We plot its values w on the w-sphere of diameter 1 into which the w-plane passes by stereographic projection. Let us denote by

$$d\tau_w = \frac{dw\,\overline{dw}}{(1 + w\overline{w})^2}$$

its surface element. $w(\mathfrak{p})$ maps the surface $\mathfrak{F}$ in a one-to-one fashion upon a Riemann surface $\mathfrak{F}_w$ covering the w-sphere. Consider now the integral

$$\frac{1}{2\pi}\int_{\mathfrak{F}} \Phi(\mathfrak{p})\,\frac{|\,w'\,|^2}{(1 + w\overline{w})^2}\,dt\,\overline{dt} \tag{6.1}$$

extended over the whole surface $\mathfrak{F}$, where t is the local uniformizing-parameter and the prime indicates differentiation with respect to t. The differential $|\,w'\,|^2\,dt\,\overline{dt}$ is independent of the particular choice of t. Assume that on the w-sphere there is a uniform charge of density 1 and that each point on $\mathfrak{F}_w$ carries the charge of its trace on the w-sphere. Then, if $\mathfrak{q}$ is the image on $\mathfrak{F}_w$ of $\mathfrak{p}$ on $\mathfrak{F}$,

$$\Phi(\mathfrak{p})\,\frac{|\,w'\,|^2}{(1 + w\overline{w})^2}\,dt\,\overline{dt}$$

is the work required to transport the charge of the surface element of $\mathfrak{F}_w$ at $\mathfrak{q}$ along any path on $\mathfrak{F}$ from infinity into the point $\mathfrak{p}$ against the potential Φ. The integral (6.1) is therefore $1/2\pi$ times the work required to assemble the charge of $\mathfrak{F}_w$ on $\mathfrak{F}$ in such a way that each point carries the charge of its image. On the other hand

$$\sum \Phi(\mathfrak{p})h(\mathfrak{p};\, \alpha_i w_2 - \beta_i w_1)\,d\tau_w$$

the sum extending over all points $\mathfrak{p}$ of $\mathfrak{F}$ is the work required to assemble in the proper places on $\mathfrak{F}$ the charges carried by the elements of $\mathfrak{F}_w$ above the surface element $d\tau_w$ of the w-sphere at the point $w = w_1/w_2$. Hence

$$\frac{1}{2\pi}\int_{\mathfrak{F}} \Phi(\mathfrak{p})\,\frac{|\,w'\,|^2}{(1 + w\overline{w})^2}\,dt\,\overline{dt} = \frac{1}{2\pi}\int \sum \Phi(\mathfrak{p})h(\mathfrak{p};\, \alpha_i w_2 - \beta_i w_1)\,d\tau_w$$

where the second integral is to be extended over the whole w-sphere. Following the argument previously employed[7] we take the average $\mathfrak{M}_{\alpha\beta}$[8] on both sides and replace w and w' by their expressions in terms of the $x_i(\mathfrak{p})$. We then obtain

$$\text{(6.2)} \qquad \frac{1}{2\pi}\int_{\mathfrak{F}} \Phi(\mathfrak{p})\mathfrak{M}_{\alpha\beta}(Q^2)\,dt\,\overline{dt} = \frac{1}{2\pi}\int \mathfrak{M}_{\alpha\beta}N(G;\alpha_i w_2 - \beta_i w_1)\,d\tau_w$$

with

$$Q = \frac{\left|\sum_{i<j}[\alpha\beta]_{ij}[xx']_{ij}\right|}{\left|\sum_0^k \alpha_i x_i\right|^2 + \left|\sum_0^k \beta_i x_i\right|^2}.$$

Next we normalize the homogeneous form w_1/w_2 of w so that

$$\bar{w}_1 w_1 + \bar{w}_2 w_2 = 1.$$

Consequently the vectors

$$\xi_i = \alpha_i w_2 - \beta_i w_1, \qquad \eta_i = \alpha_i \bar{w}_1 + \beta_i \bar{w}_2,$$

form a unitary pair if α and β do, and in particular we have

$$\sum_0^k |\xi_i|^2 = \bar{w}_1 w_1 + \bar{w}_2 w_2 = 1.$$

Hence replacing α and β in the second integral by ξ and η it becomes by Lemma 1. M.C. pg. 519

$$\frac{1}{2\pi}\int \mathfrak{M}_\xi N(G;\xi)\,d\tau_w = \tfrac{1}{2}T(G)$$

and therefore

$$\frac{1}{2\pi}\int_{\mathfrak{F}} \Phi(\mathfrak{p})\mathfrak{M}_{\alpha\beta}(Q^2)\,dt\,\overline{dt} = \tfrac{1}{2}T(G).$$

If we introduce upon the w-sphere a charge of density $\mu_w > 0$ in each point, varying with w, then the charge carried by an element of surface $d\tau_w$ will be $\mu d\tau_w$. The density μ may also depend on α and β, but if it does we presuppose that it be a homogeneous function of the combinations ξ_i. Under these conditions (6.2) is replaced by

$$\text{(6.3)} \qquad \frac{1}{2\pi}\int_{\mathfrak{F}} \Phi(\mathfrak{p})\mathfrak{M}_{\alpha\beta}\{Q^2\mu(\Xi_i)\}\,dt\,\overline{dt} = \tfrac{1}{2}\mathfrak{M}_\xi\{N(G;\xi)\mu(\xi)\}$$

where the Ξ_i arise from $\xi_i = \alpha_i w_2 - \beta_i w_1$ if we replace w_1 and w_2 by their values $\sum_0^k \alpha_j x_j(\mathfrak{p})$ and $\sum_0^k \beta_j x_j(\mathfrak{p})$ respectively:

$$\Xi_i = \sum_j (\alpha_i\beta_j - \alpha_j\beta_i)x_j(\mathfrak{p}).$$

---

[7] M C. pp. 528–530.

[8] M C. p. 518.

The value of either integral equals again the average amount of work required to assemble the total charge of an $\mathfrak{F}_w$ on $\mathfrak{F}$ in such a way that each point carries the charge of its image. With the aid of

$$N(G; \xi) \leqq T(G) + m^0(G; \xi)$$

we obtain from (6.3)

$$\frac{1}{2\pi} \int_{\mathfrak{F}} \Phi(\mathfrak{p}) \mathfrak{M}_{\alpha\beta}\{Q^2 \mu(\Xi_i)\}\, dt\, d\bar{t} \leqq cT(G) + c'(G) \tag{6.4}$$

where

$$c = \mathfrak{M}_\xi \mu(\xi), \qquad c' = \mathfrak{M}_\xi\{m^0(G; \xi)\mu(\xi)\}.$$

Again the difficulty arises that in general c' will depend on G. But c' is

$$c' = \int_{\Gamma_0} \mathfrak{M}_\xi\{\mu(\xi) \log \| \xi x \|^{-1}\}\, d\sigma$$

and therefore bounded if the mean value appearing under the integral sign is a continuous function of $x_0, x_1, \cdots, x_k$. (In the cases of meromorphic, algebroid, and ring-meromorphic curves this difficulty does not exist: c' is finite with c.)

Let us choose the density of charge on the w-sphere in a fashion which closely resembles the customary choices made in the theory of meromorphic functions.

$$\mu(\xi) = \prod_{\nu=1}^{q} \| a^{(\nu)} \xi \|^{-2\lambda_\nu} \tag{6.5}$$

where the product extends over q arbitrary points $a^{(1)}, a^{(2)}, \cdots, a^{(q)}$ in $\mathfrak{R}$. The exponents, which are assumed to be real non-negative numbers, must be determined such that $M_\xi\mu(\xi)$ is finite. With such a determination of the λ_ν the particular form (6.5) of $\mu(\xi)$ will also assure the finiteness of c'. Independently of the curve $\mathfrak{C}$ we have the relation

$$\mathfrak{M}_{\alpha\beta}\{\log Q^2 + \log \mu(\Xi_i)\} = \text{const.} + 2 \log \{X_2/X_1\} + 2 \textstyle\sum_a \lambda_a \log [\![ax]\!]^{-1} \tag{6.6}$$

for the choice (6.5) of $\mu(\xi)$.[9]

7. The Conditions on the λ_ν

Let us denote by $d\omega_\xi$ the volume element of the unitary k-sphere $\mathfrak{S}_k : \rho^2 = \sum_0^k |\xi_i|^2 = 1$. Then our task is to determine the exponents or—as we shall call them—weights λ_ν in such a manner that the integral

$$J = \int_{\mathfrak{S}_k} \mu(\xi)\, d\omega_\xi$$

[9] For a proof see: M C. pp. 530, 535.

converges. This convergence does not depend upon the nature of the curve $\mathfrak{C}$ but merely on the linear dependence scheme of the arbitrary points $a^{(\nu)}$ in $\mathfrak{R}$. It can be shown by the methods employed before (M. C. pg. 536) that J is finite for $\lambda_\nu < 1$, $(\nu = 1, 2, \cdots, q)$ if the points $a^{(\nu)}$ satisfy no accidental linear relations, i.e. if any $(k + 1)$ of them are linearly independent. From this restriction we wish to free ourselves.

The integral whose convergence we are to investigate can be written in the form

$$J = \int d\omega_\xi/\Pi(\xi), \qquad \Pi(\xi) = \prod_{\nu=1}^{q} |(a^{(\nu)}\xi)|^{2\lambda_\nu}$$

where the product $\Pi(\xi)$ is formed over a finite number of linear forms $(a\xi)$ whose coefficients a_i are normalized so as to satisfy the condition $\sum_0^k |a_i|^2 = 1$. Each of them is furthermore provided with a weight $\lambda \geqq 0$.

We interprete the ξ_i as components of a vector in a complex $(k + 1)$-dimensional vector space. The sum $\sum \lambda$ extended over all planes $(a\xi)$ which contain a given l-dimensional linear submanifold S_l of that vector space shall be called the load carried by S_l.

First we investigate the somewhat simpler space integral

$$\int dV_\xi/\Pi(\xi) \qquad \text{extended over } \rho^2 \leqq r_0^2. \tag{7.1}$$

Before proceeding any further it is necessary to express the volume element dV_ξ of the real $(2k + 2)$-dimensional space with complex coordinates ξ_i. What we are seeking is the generalization of the surface element $dz\,d\bar{z}$ employed instead of $dx\,dy$ when the (x, y)-plane is described by the complex coordinate $z = x + iy$.

Let our space be spanned by the $2k + 2$ real basis vectors

$$\mathfrak{e}_0, \mathfrak{e}_0', \mathfrak{e}_1, \mathfrak{e}_1', \cdots, \mathfrak{e}_k, \mathfrak{e}_k'.$$

We split our coordinates ξ_i into real and imaginary parts

$$\xi = x + iy \qquad (i = \sqrt{-1})$$

as suggested by the two-dimensional case. An arbitrary line element

$$\mathfrak{d} = (d\xi_0, d\xi_1, \cdots, d\xi_k)$$

then has the components

$$dx_0, dy_0, dx_1, dy_1, \cdots, dx_k, dy_k$$

with respect to the above vector basis. Let the line element

$$(id\xi_0, id\xi_1, \cdots, id\xi_k)$$

with the corresponding complex components be called $\mathfrak{d}'$. We note that the transition from $\mathfrak{d}$ to $\mathfrak{d}'$ is invariant under an analytic coordinate transformation. The real components of $\mathfrak{d}'$ are

$$-dy_0, dx_0, -dy_1, dx_1, \cdots, -dy_k, dx_k.$$

If we have $(k + 1)$ line elements

$$\mathfrak{d}_0 = (d\xi_0, d\xi_1, \cdots, d\xi_k)$$
$$\mathfrak{d}_1 = (\delta\xi_0, \delta\xi_1, \cdots, \delta\xi_k)$$
$$\vdots$$

linearly independent in the complex sense then $\mathfrak{d}_0, \mathfrak{d}_0', \cdots, \mathfrak{d}_k, \mathfrak{d}_k'$, are linearly independent in the real sense. As volume element we use the parallelopiped spanned by them:

$$dV_\xi = \begin{vmatrix} dx_0 & dy_0 & dx_1 & dy_1 & \cdots \\ -dy_0 & dx_0 & -dy_1 & dx_1 & \cdots \\ \delta x_0 & \delta y_0 & \delta x_1 & & \\ \vdots & \vdots & \vdots & & \end{vmatrix}.$$

From their definitions it follows that

$$\mathfrak{d} + i\mathfrak{d}' = \overline{d\xi}_0(\mathfrak{e}_0 + i\mathfrak{e}_0') + \cdots + \overline{d\xi}_k(\mathfrak{e}_k + i\mathfrak{e}_k'),$$
$$\mathfrak{d} - i\mathfrak{d}' = d\xi_0(\mathfrak{e}_0 - i\mathfrak{e}_0') + \cdots + d\xi_k(\mathfrak{e}_k - i\mathfrak{e}_k'),$$

hence the desired volume element appears in the form

$$dV_\xi = \bar{\Delta}\Delta$$

where

$$\Delta = \begin{vmatrix} d\xi_0 & d\xi_1 & \cdots & d\xi_k \\ \delta\xi_0 & \delta\xi_1 & \cdots & \delta\xi_k \\ \vdots & \vdots & & \vdots \end{vmatrix}$$

Lemma 1. **The integral (7.1) is convergent if each linear vector manifold of dimension $(k + 1 - l)$ carries a load**

$$\Lambda_l < l. \qquad (l = 1, 2, \cdots, k + 1).$$

The proof shall be carried through by means of an induction with respect to the number of dimensions.

Let us assume that among the planes $(a^{(\nu)}\xi) = 0$ there are at least $(k + 1)$ linearly independent ones. Should this not be the case we can introduce additional linear forms with the weights $\lambda = 0$.

The set $\mathfrak{A}$ of planes a: $(a\xi) = 0$, $\sum_0^k | a_i |^2 = 1$, in the $(k + 1)$ -dimensional vector space $\mathfrak{E}_{k+1}$ shall shortly be referred to as the configuration $\{a\}$. The intersection of any k linearly independent planes $(a^{(\nu)}\xi) = 0$ shall be called a vertex $\mathfrak{z}$. It is our intention to find corresponding to every configuration of this kind a cell-division of the ξ-space such that each vertex $\mathfrak{z}$ is contained in a cell $\mathfrak{Z}$ which is cut by no other planes of $\mathfrak{A}$ except those which pass through $\mathfrak{z}$. This can be accomplished inductively, for any plane a of $\mathfrak{A}$ is an $\mathfrak{E}_k$ and its in-

tersections $\bar{b}$ with all planes of $\mathfrak{A}$ different from a form a configuration in this lower-dimensional space. Assume therefore that we have found a cell-division with the desired property for the configuration $\{\bar{b}\}$ in this $\mathfrak{E}_k$. Let $\mathfrak{Z}_k$ be one of its cells with vertex $\mathfrak{z}$, a flat cell, so to speak, which we have to extend into space. To this end we perform a unitary transformation

$$(\xi_0, \xi_1, \cdots, \xi_k) \to (\zeta_0, \zeta_1, \cdots, \zeta_k)$$

such that $(a\xi) = \zeta_0$, i.e. such that the ζ_0-axis is orthogonal to the plane $(a\xi) = 0$. Then we define

$$(\zeta_0, \zeta_1, \cdots, \zeta_k) \text{ is in } \mathfrak{Z}_{k+1} \text{ if}$$

$$\text{D 1.)} \quad (0, \zeta_1, \zeta_2, \cdots, \zeta_k) \text{ is in } \mathfrak{Z}_k.$$

$$\text{D 2.)} \quad |\zeta_0| \leqq |(b\zeta)| \text{ for all } b \text{ in } \mathfrak{A}.$$

This inductive definition begins with

$$\mathfrak{z} = \mathfrak{Z}_1.$$

Let us call $\mathfrak{z}$ the *center* and $\zeta_0 = 0$ the *base* of $\mathfrak{Z}_{k+1}$. The cells thus defined have the following properties:

(1) They are cones in the sense that $(\zeta_0, \zeta_1, \cdots, \zeta_k)$ in $\mathfrak{Z}_{k+1}$ implies $(\lambda\zeta_0, \cdots, \lambda\zeta_k)$ in $\mathfrak{Z}_{k+1}$ for any complex constant λ.

(2) They are closed, as evidenced by the equality sign in D 2.).

(3) They cover the whole space: Let $\zeta^{(0)}$ be any point of our space then there will be an $|(a\zeta^{(0)})|$ such that none of the expressions $|(b\zeta^{(0)})|$ (b in $\mathfrak{A}$, $b \neq a$) have a smaller value, consequently $\zeta^{(0)}$ belongs to a cell with the base a.

(4) A cell contains but one vertex $\mathfrak{z}$, namely its center.

If $\mathfrak{z}'$ were in $\mathfrak{Z}$ then the k linearly independent planes defining $\mathfrak{z}'$ would also have to go through $\mathfrak{z}$ which is impossible unless $\mathfrak{z}' \equiv \mathfrak{z}$. Finally we have

(5) A plane b of $\mathfrak{A}$ has no point other than the origin $(0, 0, \cdots, 0)$ in common with a cell unless it passes through its center.

Assume this to be true for the flat cells $\mathfrak{Z}_k$ in $\mathfrak{E}_k$ and the manifolds $\bar{b}$ of the configuration $\{\bar{b}\}$. Let b be a plane such that for some point $\zeta^{(0)} \neq (0, \cdots, 0)$ in $\mathfrak{Z}_{k+1}$ we have $(b\zeta^{(0)}) = 0$. Either b is the plane $\zeta_0 = 0$ or it is not. In either case b is seen to pass through $\mathfrak{z}$. In the first one this is evident; in the second one b intersects the plane $\zeta_0 = 0$ in an $\mathfrak{E}_{k-1}$ shortly denoted by $\bar{b}$. We remark that it follows from $(b\zeta^{(0)}) = 0$ that also $\zeta_0^{(0)} = 0$ on account of D 2.). Hence the point $\zeta^{(0)}$ is firstly in the flat cell $\mathfrak{Z}_k$ and secondly contained in $\bar{b}$; therefore it follows from the induction hypothesis that $\bar{b}$ must contain $\mathfrak{z}$, implying the same for b. The contention is evidently true for $k = 1$ since every $\mathfrak{E}_1$ of our configuration is some cell's center, any two of them co-inciding completely as soon as they have a point other than $(0, 0, \cdots, 0)$ in common.

For any plane b of $\mathfrak{A}$ not passing through $\mathfrak{z}$ we even have the quantitative estimate

$$|(b\zeta)| \geqq \delta\left(\textstyle\sum_0^k |\zeta_i|^2\right)^{\frac{1}{2}} \tag{7.2}$$

with a positive constant δ for all points ζ of a cell $\mathfrak{Z}$ with center $\mathfrak{z}$. This follows at once from the fact that the function $\left(\sum_0^k |\zeta_i|^2\right)^{\frac{1}{2}} : |(b\zeta)|$ is a bounded function of its arguments on the closed set $\mathfrak{Z}$. The quantity

$$\beta = |(b\mathfrak{z}_i^{(0)})| : \left(\textstyle\sum_0^k |\mathfrak{z}_i^{(0)}|^2\right)^{\frac{1}{2}}$$

has the same value for all points $\mathfrak{z}^{(0)}$ of $\mathfrak{z}$ (The "distance" of the vertex $\mathfrak{z}$ from the plane b). If b does not pass through $\mathfrak{z}$ we have $\beta > 0$ and can derive the explicit estimate

$$\delta \geqq 2^{-k}\beta. \qquad \text{(See Appendix)}$$

If we unite all cells $\mathfrak{Z}_{k+1}$ with center $\mathfrak{z}$ we obtain the star $\mathfrak{Z}_{k+1}^*$. All points of $\mathfrak{z}$ with exception of $(0, 0, \cdots, 0)$ will be interior points of $\mathfrak{Z}_{k+1}^*$. The estimate (7.2) remains true for all points of $\mathfrak{Z}_{k+1}^*$ if I choose for δ the smallest one of all those δ's which correspond to the cells combined into $\mathfrak{Z}_{k+1}^*$.

We designate by $\mathfrak{Z}^{(e)}$ the intersection of $\sum_0^k |\zeta_i|^2 \leqq r_0^2$ and $\mathfrak{Z}_{k+1}^* : \mathfrak{Z}^{(e)} = \{\sum_0^k |\zeta_i|^2 \leqq r_0^2\} \frown \mathfrak{Z}_{k+1}^*$. The region over which (7.1) is to be integrated consists of a finite number of such stumps $\mathfrak{Z}^{(e)}$. In order to evaluate the part of (7.1) extending over any one of them we shall specify the coordinate system $(\zeta_0, \zeta_1, \cdots, \zeta_k)$ more accurately than has been done so far. We choose a cell $\mathfrak{Z}_{k+1}$ of the star $\mathfrak{Z}_{k+1}^*$: Let its base be $\mathfrak{E}_k$. The corresponding flat cell will again have a base $\mathfrak{E}_{k-1}$, and so forth. Now we determine the $\zeta_0, \zeta_1, \cdots, \zeta_k$ such that

$$\begin{aligned}
&\mathfrak{E}_k \quad \text{is } \zeta_0 = 0,\\
&\mathfrak{E}_{k-1} \text{ is } \zeta_0 = 0,\ \zeta_1 = 0,\\
&\qquad\vdots\\
\mathfrak{z} = {}&\mathfrak{E}_1 \quad \text{is } \zeta_0 = 0,\ \zeta_1 = 0, \cdots, \zeta_{k-1} = 0.
\end{aligned}$$

We demand furthermore that the ζ_i arise from the ξ_i by a unitary transformation. This determines them uniquely to within factors of absolute value 1. We refer to the system of the ζ_i-axes as the *frame work* spanning the cell $\mathfrak{Z}_{k+1}$. The coordinate system which we introduce in the star $\mathfrak{Z}_{k+1}^*$ is the frame work of any one of its cells.

Let $(b\zeta) = 0$ be some plane b of $\mathfrak{A}$ not going through $\mathfrak{z}$. Then the planes $\zeta_0 = 0, \zeta_1 = 0, \cdots, \zeta_{k-1} = 0, (b\zeta) = 0$, form a system of $(k + 1)$ linearly independent planes in terms of which any plane of our configuration $\{a\}$ can be expressed in linear fashion. We denote by $c_0, c_1, \cdots, c_{k-1}$, the coefficients of $\zeta_0, \zeta_1, \cdots, \zeta_{k-1}$, and by c_k the coefficient of $(b\zeta)$ in this representation:

$$(a\zeta) \equiv \textstyle\sum_0^{k-1} c_i\zeta_i + c_k(b\zeta).$$

Furthermore we have from (7.2) that $|b_k| \geqq \delta$.

Now we perform the substitution

$$\begin{aligned}(b\zeta) &= z\\ \zeta_0 &= \eta_0 z\\ \zeta_1 &= \eta_1 z\\ &\vdots\\ \zeta_{k-1} &= \eta_{k-1} z\,.\end{aligned}$$

The successive transformations

$$(\xi_0, \xi_1, \cdots, \xi_k) \rightarrow (\zeta_0, \zeta_1, \cdots, \zeta_{k-1}, (b\zeta)) \rightarrow (z, \eta_0, \cdots, \eta_{k-1})$$

bring our volume element into the form

$$dV_\xi = |b_k|^{-2}\, dV_\zeta = |b_k|^{-2}(z\bar{z})^k\, dz\, d\bar{z}\, dV_\eta$$

and our integral can now be written as

$$\int dV_\xi/\Pi(\xi) = |b_k|^{-2}\int (z\bar{z})^{k-\Lambda}\, dz\, d\bar{z}\cdot \prod_1^q |c_0^{(\nu)}\eta_0 + \cdots + c_{k-1}^{(\nu)}\eta_{k-1} + c_k^{(\nu)}|^{-2\lambda_\nu}\, dV_\eta$$

where $\Lambda = \sum_1^q \lambda_\nu$ is the total load carried by the origin. Both integrations are to be extended over the stump $\mathfrak{Z}^{(s)}$.

The linear forms in the denominator of the second factor will be divided into two classes—the first one containing the forms corresponding to planes that do—the second one those corresponding to planes that do not pass through $\mathfrak{z}$. For any one of the latter kind we have

$$\left| \sum_0^{k-1} c_i\zeta_i + c_k(b\zeta) \right| \geqq \delta'\left\{\sum_0^k |\zeta_i|^2\right\}^{\frac{1}{2}} \geqq \delta' |z| \qquad (c_k \neq 0)$$

hence

$$|c_0\eta_0 + \cdots + c_{k-1}\eta_{k-1} + c_k| \geqq \delta' \qquad (\delta' = \text{const.} > 0).$$

This permits us to write

$$\int dV_\xi/\Pi(\xi) \leqq c\int (z\bar{z})^{k-\Lambda}\, dz\, d\bar{z}\cdot \prod_\mu |c_0^{(\mu)}\eta_0 + \cdots + c_{k-1}^{(\mu)}\eta_{k-1}|^{-2\lambda_\mu}\, dV_\eta \tag{7.3}$$

where $c > 0$ is some constant, and the product in the denominator of the second term is to be extended over all those planes $a^{(\mu)}$ of $\mathfrak{A}$ which pass through $\mathfrak{z}$. Again both integrals are to be extended over $\mathfrak{Z}^{(s)}$.

Concerning the limits of integration in the new variables we note that from $\sum_0^k |b_i|^2 = 1$ and $\sum_0^k |\zeta_i|^2 \leqq r_0^2$ follows $|z| \leqq r_0$ since

$$\frac{|(b\zeta)|^2}{\sum_0^k |\zeta_i|^2} = \frac{|z|^2}{\sum_0^k |\zeta_i|^2} \leqq 1.$$

On the other hand we have from (7.2)

$$|z| \geqq \delta\left\{\sum_0^k |\zeta_i|^2\right\}^{\frac{1}{2}} \geqq \delta\left\{\sum_0^{k-1} |\zeta_i|^2\right\}^{\frac{1}{2}},$$

finally yielding

$$| \eta_0 |^2 + | \eta_1 |^2 + \cdots + | \eta_{k-1} |^2 \leqq \delta^{-2}.$$

In other words: The stump $\mathfrak{Z}^{(s)}$ is contained in the region described by

$$| z | \leqq r_0, \quad \sum_0^{k-1} | \eta_i |^2 \leqq \delta^{-2}.$$

Integrating (7.3) over this region rather than $\mathfrak{Z}^{(s)}$ will increase the integral. We introduce polar coordinates for the complex variable

$$z = \rho e^{i\varphi} \qquad (i = \sqrt{-1})$$

and thus obtain

$$\int_{\mathfrak{Z}^{(s)}} \frac{dV_\xi}{\Pi(\xi)} \leqq c \int_0^{r_0} \int_0^{2\pi} \rho^{2k+1-2\Lambda} \, d\rho \, d\varphi \cdot \int \frac{dV_\eta}{\Pi(\eta)}$$

where the last integral is to be extended over the sphere

$$\sum_0^{k-1} | \eta_i |^2 \leqq \delta^{-2}.$$

The convergence of the first factor is guaranteed by the condition $\Lambda < k + 1$. The convergence of the second factor is the contention of Lemma 1 with k replaced by $(k - 1)$, i.e. the induction hypothesis.

We anchor this induction at $k = 1$ for which the lemma is evidently true.

The integral J converges under the reduced assumption which arises from the one of Lemma 1 by dropping the restriction that the origin carries a load $< k + 1$. This is shown by modifying the first step of our proof, integrating (7.1) over a spherical shell

$$0 < r_1 \leqq \rho \leqq r_0 \tag{7.4}$$

rather than over a solid sphere.

$$\int dV_\xi / \Pi(\xi) = J \int_{r_1}^{r_0} \rho^{2k+1-2\Lambda} \, d\rho,$$

where the first integral is extended over the shell (7.4). Concerning the limits of integration we note that the intersection of $\mathfrak{Z}_{k+1}^*$ with the shell is contained in the region

$$\delta r_1 \leqq | z | \leqq r_0, \quad \sum_0^{k-1} | \eta_i |^2 \leqq \delta^{-2}.$$

Therefore the part of J extending over the intersection of $\mathfrak{Z}_{k+1}^*$ with the unitary k-sphere $\mathfrak{S}_k$ will be less than or equal to

$$\text{const} \int dV_\eta / \Pi(\eta) \text{ extended over } \sum_0^{k-1} | \eta_i |^2 \leqq \delta^{-2}.$$

Thus the proof of what we maintained is reduced to an application of Lemma 1 for $(k - 1)$ instead of k. Returning to the projective way of expression and considering the ξ_i as homogeneous plane-coordinates, the $a_i^{(\nu)}$ as homogeneous point-coordinates, we obtain the

LEMMA 2. The integral J converges if each $(l-1)$-dimensional linear subspace of the projective k-space $(1 \leqq l \leqq k)$ carries a load $\Lambda_l < l$. (The load Λ_l of a given subspace is the sum of the weights of the points $a^{(\nu)}$ contained therein.)

The proof of this lemma indicates the reason why no restriction need be made upon the total load carried by the whole space. It is the load of the origin in the ξ-space and of no influence upon the convergence of the integral J since the latter extends only over points $\sum_0^k |\xi_i|^2 = 1$.

For the choice (6.5) of $\mu(\xi)$ the average $\mathfrak{M}_\xi \mu(\xi)$ will remain finite if the weights λ_ν are so chosen that any given $(l-1)$-spread in $\mathfrak{R}$ $(l = 1, \cdots, k)$ carries a load $\Lambda_l < l$. This, as we remarked once before, will imply the finiteness of

$$c' = \mathfrak{M}_\xi\{m^0(G;\xi)\mu(\xi)\}$$

as well. Let us assume therefore that in the following special cases the λ_ν always comply with the conditions of Lemma 2.

8. The Third Main Theorem for Algebroid Curves

Once more we turn to the case where $\mathfrak{F}$ is an n-sheeted unbounded covering surface of the finite z-plane. We take the set of $(k+1)$ functions defining a realization of $\mathfrak{F}$ and substitute them for the x_i in (6.6). The prime in

$$X_2^2 = \sum_{i<j} |[xx']_{ij}|^2$$

we interpret as differentiation with respect to z. Multiplying through by $d\sigma = (1/2\pi)\,d\varphi$ and integrating over the boundary of G we shall obtain, if we make use of the concavity of the logarithmic function,

$$2\Omega_1(r) + 2\sum_a \lambda_a m^*(r;a) + \text{const} \leqq \log\left\{\frac{1}{2\pi n}\int_0^{2\pi n} \mathfrak{M}_{\alpha\beta}\{Q^2\mu(\Xi_i)\}\,d\varphi\right\} = \Theta(r).$$

We indicate the rotational symmetry of the potential Φ by writing

$$\Phi(z) = \Phi(\rho) \quad \text{when} \quad |z| = \rho.$$

Then (6.3) yields for $\Theta(\rho)$ the relation

$$\int_0^r \Phi(\rho)e^{\Theta(\rho)}\rho\,d\rho = \int_{r_0}^r \frac{dr}{r}\int_0^r e^{\Theta(\rho)}\rho\,d\rho \leqq cT(r) + c',$$

denoted shortly by

$$\Theta(r) = \omega(T(r)).$$

From this it follows in customary fashion that

$$\Theta(r) < \kappa \log T(r) - 2\log r,$$

an inequality holding "almost everywhere," i.e. outside of certain intervals I_r

whose logarithmic measure is finite: $\int_{I_r} \frac{dr}{r} < \infty$.[10] κ is an arbitrary constant $\kappa > 1$, a meaning which this symbol shall retain throughout this chapter.

Formulating the resulting relations at once for higher l we have

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) = \tfrac{1}{2}\omega(T_l(r))$$

or, in the form of an inequality holding almost everywhere,

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) < \frac{\kappa}{2} \log T_l(r) - \log r$$

with weights λ_a attached to the points a and so chosen that the total load carried by any one $(h - 1)$-spread $(1 \leqq h \leqq k_l)$ in $\Re_l$ is less than h.

These so-called defect relations constitute the third main theorem, holding in this form regardless of any accidental linear relations between the points over which the sum on the left is to be extended, i.e. in particular those points for which m^* is not ~ 0. (Exceptional points)

The methods[11] used for the derivation of estimates of the subsequent nature yield again

$$\Omega_l(r) < \frac{\kappa}{2} \log T(r)$$

hence

$$T_l(r) < lT(r) + \kappa \frac{l(l+1)}{4} \log T(r)$$

and, introducing the symbols

$$\Delta_l(r) = V_l(r) + \sum_a \lambda_a m_l^*(r; a),$$

we finally obtain

$$\Delta_l(r) < \frac{k+1}{k+1-l} T(r) + \frac{k(k+1)}{2(k+1-l)} J(r) + \kappa \frac{k(k+1)}{4(k+1-l)} \log T(r) \tag{8.1}$$

holding almost everywhere. To complete our results we have to give an estimate for the density $J(r)$ of branchpoints of $\mathfrak{F}$ over the z-plane. E. Ullrich proves[12] that for the characteristic $T^*(r)$ of any algebroid function $w = f(z)$, one-valued on $\mathfrak{F}$,

$$J(r) \leqq 2(n-1)T^*(r).$$

This characteristic $T^*(r)$, as defined by E. Ullrich, is in the sense of equivalence equal to the order of the algebroid curve $w_1/w_2 = w$ defined in the complex

[10] M C. pp. 527–528.
[11] M C. pp. 532–533.
[12] l. c. pp. 209–210.

one-dimensional projective space $\{w_1, w_2\}$ as a realization of $\mathfrak{F}$. Such a curve will certainly be defined by projecting our curve $\mathfrak{C}$ from some $(k-2)$-dimensional linear subspace of $\mathfrak{R}$. Application of (3.5) shows us that

$$T^*(r) \leqq T(r) + \text{const.}$$

since in our case

$$\int_{\Gamma_0} \log (X : X')\, d\sigma = \text{const.}$$

The resulting relation

$$J(r) \leqq 2(n-1)T(r) + \text{const.} \tag{8.2}$$

shows that not only the level of transcendency but also the degree of ramification of an algebroid curve is set by the first order $T(r)$. By means of (8.2) we obtain from (8.1) the relations

$$V_l(r) + \sum_a \lambda_a m_l^*(r; a) < \frac{(k+1)(nk-k+1)}{k+1-l}\, T(r) + S(r)$$

holding almost everywhere with

$$S(r) = O\,(\log rT(r)).$$

Thus the defect relation appear in the form that was given them by R. Nevanlinna: $k = 1$; $n = 1$, and E. Ullrich: $k = 1$.

9. The Third Main Theorem for Ring-meromorphic Curves

Instead of the function $x_i(\mathfrak{p})$, meromorphic on the doubly punctured plane $\mathfrak{F}$, which define its realization $\mathfrak{C}$ we use the functions $x_i(z)$ defined for z ranging over that part of the z-plane which is covered by $\mathfrak{F}$. For the choice of G_0 and the exhausting sequence $\{G_{R,r}\}$ made heretofore the potential $\Phi(z)$ is a function of $|z| = \rho$ rather than of z:

$$\Phi_G(z) = \Phi(R, r; \rho).$$

From (6.6) it follows that for

$$\Theta(\rho) = \log \left\{\frac{1}{2\pi}\int_0^{2\pi} \mathfrak{M}_{\alpha\beta}\{Q^2\mu(\Xi_i(\rho e^{i\varphi}))\}\, d\varphi\right\}$$

we have

$$2\Omega_1(\rho) + 2\sum_a \lambda_a m^*(\rho; a) + \text{const} \leqq \Theta(\rho)$$

or

$$2\Omega_1(G_{R,r}) + 2\sum_a \lambda_a m^*(G_{R,r}; a) + \text{const} \leqq [\Theta(R); \Theta(r)].$$

On the other hand (6.3) shows that

$$\int_r^R \Phi(R, r; \rho) e^{\Theta(\rho)} \rho\, d\rho \leqq cT(G_{R,r}) + c'$$

with constants c and c'. To set in evidence the fact that the integral on the left also possesses the characteristic pattern of all quantities related to ring-meromorphic curves we write it as

$$\int_r^R \Phi(R, r; \rho)e^{\Theta(\rho)}\rho\, d\rho = [{}^e\Psi(R); {}^i\Psi(r)]$$

with

$${}^e\Psi(R) = \int_{R_0}^R \frac{dR}{R}\int_{R_0}^R e^{\Theta(\rho)}\rho\, d\rho + \log\frac{R}{R_0}\left\{\int_{r_m}^{R_0} e^{\Theta(\rho)}\rho\, d\rho\right\},$$

$${}^i\Psi(r) = \int_r^{r_0} \frac{dr}{r}\int_r^{r_0} e^{\Theta(\rho)}\rho\, d\rho + \log\frac{r_0}{r}\left\{\int_{r_0}^{r_m} e^{\Theta(\rho)}\rho\, d\rho\right\}.$$

Thus we finally obtain

$$[{}^e\Psi(R); {}^i\Psi(r)] \leqq c[{}^eT(R); {}^iT(r)] + c'. \tag{9.1}$$

The arbitrariness

$${}^e\Psi(R) \to {}^e\Psi(R) + c\log\frac{R}{R_0}, \quad {}^i\Psi(r) \to {}^i\Psi(r) - c\log\frac{r_0}{r}$$

finds again an immediate geometric expression in the freedom of choice of the intermediate circle of radius r_m.

Suppose two functions $\Theta(r)$ and $T(r)$ are defined for sufficiently large values of r. Then we agreed[13] to write

$$\Theta(r) = \omega(T(r))$$

if there exists a centrally symmetric potential $U(r)$ due to a distribution of free charge of density $\exp\Theta(r)$:

$$\frac{1}{r}\frac{\partial}{\partial r}\left\{r\frac{\partial U(r)}{\partial r}\right\} = \exp\Theta(r),$$

such that

$$U(r) \leqq cT(r). \qquad (c = \text{const.} > 0)$$

It follows that there exist constants of integration a and b such that

$$U(r) = \int_{r_0}^r \frac{dr}{r}\int_{r_0}^r \exp\Theta(\rho)\rho\, d\rho - a\log r - b$$

satisfies this relation, r_0 being an arbitrary lower bound for the integrations, subject only to the condition that $\Theta(r)$ and $T(r)$ are defined for $r \geqq r_0$.

We note concerning the relation $\Theta = \omega(T)$ that it implies:

(1) $\Theta(r) = \omega(T^*(r))$ for any $T^*(r) = T(r) + c'\log r + c''$ with constants c' and c'', because the additional terms correspond to free charges of density zero.

[13] M C. p. 528.

(2) $\Theta'(r) = \omega(T)$ for any $\Theta' \leqq \Theta$, because

$$U'(r) = \int_{r_0}^{r} \frac{dr}{r} \int_{r_0}^{r} \exp \Theta'(\rho)\rho\, d\rho - a \log r - b \leqq U(r).$$

(3) The existence of constants c' and c'' such that

$$T^*(r) > 0$$

because for

$$T^*(r) = T(r) + \frac{a}{c} \log r + \frac{b}{c}$$

we have

$$cT^*(r) \geqq \int_{r_0}^{r} \frac{dr}{r} \int_{r_0}^{r} \exp \Theta(\rho)\rho\, d\rho > 0.$$

(4) $\Theta(r) < \kappa \log T^*(r)$ almost everywhere if the constants c' and c'' are so chosen as to make $T^*(r) > 0$.

The additional term of the order of $\log r$ has been neglected in the presence of the term $\log T^*$, compared to which it is of importance only in the cases of lowest transcendency.

We extend the meaning of this symbol to functions $\Theta(r)$ and $T(r)$ defined for sufficiently small values of r by writing

$$\Theta(r) = \omega(T(r))$$

if after an inversion on the unit-circle: $r \to 1/r = \tilde{r}$ we have for the functions $\tilde{\Theta}(\tilde{r})$ and $\tilde{T}(\tilde{r})$, defined by

$$\tilde{\Theta}(\tilde{r}) = \Theta(r), \qquad \tilde{T}(\tilde{r}) = T(r),$$

the relation

$$\tilde{\Theta}(\tilde{r}) = \omega(\tilde{T}(\tilde{r})).$$

The application to our case is evident since both ${}^e\Psi$ and ${}^i\Psi$ are centrally symmetric potentials, each due in its region of definition ($r > R_0$, $r < r_0$ respectively) to a distribution of free charges of density $\exp \Theta(r)$. By choosing

$$r = e^{-1}r_0\text{, while } R \text{ remains variable,}$$

or

$$R = eR_0\text{, while } r \text{ remains variable}$$

we obtain from (9.1) the inequalities

$${}^e\Psi(R) \leqq c\{{}^eT(R) + \text{const.} \log R + O(1)\}$$

$${}^i\Psi(r) \leqq c\{{}^iT(r) + \text{const.} \log r + O(1)\}$$

with

$$c = \mathfrak{M}_\xi \mu(\xi) > 0.$$

Consequently

$$\Theta(r) = \omega({}^eT(r)), \qquad \Theta(r) = \omega({}^iT(r)),$$

and therefore almost everywhere

$$\Theta(r) < \kappa \log {}^eT^*(r) \qquad \text{for } r > R_0,$$

$$\Theta(r) < \kappa \log {}^iT^*(r) \qquad \text{for } r < r_0,$$

with constants chosen such that ${}^eT^*$ and ${}^iT^*$ are > 0 (see Section 5). Thus the defect relations appear in the form

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) = \tfrac{1}{2}\omega({}^eT_l(r)) \qquad \text{for } r > R_0,$$

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) = \tfrac{1}{2}\omega({}^iT_l(r)) \qquad \text{for } r < r_0,$$

or as inequalities holding almost everywhere

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) < \frac{\kappa}{2} \log {}^eT_l^*(r) \qquad \text{for } r > R_0,$$

$$\Omega_l(r) + \sum_a \lambda_a m_l^*(r; a) < \frac{\kappa}{2} \log {}^iT_l^*(r) \qquad \text{for } r < r_0.$$

Observing that

$$[\log {}^eT^*; \log {}^iT^*] \leqq \log [{}^eT^*; {}^iT^*]$$

allows us to combine each of the above pairs into a single relation:

$$\Omega_l(G) + \sum_a \lambda_a m_l^*(G; a) = \frac{\kappa}{2} \log T_l(G). \qquad (G \equiv G_{R,r})$$

Again we might derive estimates of the kind

$$T_l < lT + S$$

$$V_l < \frac{k+1}{k+1-l}\, T + S$$

climaxing in

$$V_l + \sum_a \lambda_a m_l^* < \frac{k+1}{k+1-l}\, T + S. \tag{9.2}$$

These relations hold for the averaged quantities as well as for the unaveraged ones, if in the latter case we supplement them by the conventions governing the pair-wise connection between the unaveraged constituents. S stands throughout for a term $O(\log T)$.

Applying (9.2) to a ring-meromorphic function, for instance to one of the quotients x_i/x_0 which define the realization $\mathfrak{C}$, it is seen that any such function must assume every value with the possible exception of at most two. This is nothing new considering that every ring-meromorphic function $f(z)$ is changed by the substitution

$$z = e^{\zeta}$$

into a function meromorphic and periodic of period $2\pi i$ on the open ζ-plane. The validity of Picard's theorem for $f(\exp \zeta) = F(\zeta)$ yields an immediate proof of the above statement. If we write $\zeta = x + iy$ it follows that $F(\zeta)$ repeats its values in every infinite strip parallel to the x-axis of width 2π. Hence the surface $\mathfrak{F}$ on which these functions are defined presents itself as a circular cylinder of radius 1 around the x-axis. The order of a function on $\mathfrak{F}$ is described by a pair of functions ${}^eT(x)$ and ${}^iT(x)$:

$$T = [{}^eT(X); {}^iT(x)].$$

The first one is defined for $X > x_1$, the second one for $x < x_0$, $(x_1 > x_0)$, and they are subject to the condition

$$[{}^eT(X); {}^iT(x)] \sim [{}^eT(X) + cX; {}^iT(x) + cx].$$

Ring-meromorphic curves present themselves in this light as realizations in k-space of the cylindrical surface $\mathfrak{F}$ in abstracto.

Appendix

Suppose we are given a cell-division corresponding to a configuration $\{a\}$ of planes $a: \sum_0^k | a_i |^2 = 1$, in an $(k + 1)$-dimensional vector space $\mathfrak{E}_{k+1}$: $\{\xi_0, \xi_1, \cdots, \xi_k\}$, of the sort described in Section 7. Concerning it we stated the following estimate: If b is a plane not passing through the center $\mathfrak{z}$ of a cell $\mathfrak{Z}$ then

$$(1) \qquad | (b\xi) | : \{ \textstyle\sum_0^k | \xi_i |^2 \}^{\frac{1}{2}} \geqq 2^{-k}\beta$$

for any point ξ of $\mathfrak{Z}$, where

$$\beta = | (b\mathfrak{z}^{(0)}) | : \{ \textstyle\sum_0^k | \mathfrak{z}_i^{(0)} |^2 \}^{\frac{1}{2}},$$

$\mathfrak{z}^{(0)}$ being any point on the vertex $\mathfrak{z}$. In a conversation H. Weyl proposed to me the following proof of this inequality.

As a coordinate system in $\mathfrak{Z}$ we introduce its frame work $\{\zeta_0, \cdots, \zeta_k\}$. Let $(b\zeta) = 0$ be the equation of the plane b; then we have $\beta = | b_k | > 0$. For any point ζ for which $\sum_0^k | \zeta_i |^2 = 1$ (a restriction which does not impair the generality of our argument considering the conical shape of $\mathfrak{Z}$) we now write

$$| (b\zeta) | = s;$$

hence

$$| \zeta_0 | \leqq s \qquad \text{on account of D 2).}$$

Consequently

$$|b_1\zeta_1 + \cdots + b_k\zeta_k| \leqq s + |b_0\zeta_0| \leqq s(1 + |b_0|). \tag{2}$$

But from (7.2), applied to the space $\zeta_0 = 0$ of one dimension less, it follows that

$$\begin{aligned} |b_1\zeta_1 + \cdots + b_k\zeta_k|^2 &\geqq \delta_k^2(1 - |b_0|^2)(1 - |\zeta_0|^2) \\ &\geqq \delta_k^2(1 - |b_0|^2)(1 - s^2), \qquad (\delta_k > 0). \end{aligned} \tag{3}$$

The factor $(1 - |\zeta_0|^2)$ replaces $|\zeta_1|^2 + \cdots + |\zeta_k|^2$ which it equals since $\sum_0^k |\zeta_i|^2 = 1$, while $(1 - |b_0|^2)$ makes its appearance because (7.2) was derived under the assumption that $\sum_0^k |b_i|^2 = 1$; hence in the present case we have to multiply on the right by $|b_1|^2 + \cdots + |b_k|^2 = 1 - |b_0|^2$ in order to make the relation completely analogous. Combining (2) and (3) we obtain

$$\frac{s^2}{1 - s^2} \geqq \delta_k^2 \frac{1 - |b_0|}{1 + |b_0|}. \tag{4}$$

Hence, writing $|b_0| = \cos \alpha_k$, the right side of (4) will become $\delta_k \operatorname{tg}(\alpha_k/2)$ and the inequality states that (7.2) will be satisfied for any $\delta_{k+1} = \delta$ chosen such that

$$\frac{\delta_{k+1}^2}{1 - \delta_{k+1}^2} \leqq \delta_k^2 \operatorname{tg}^2 \frac{\alpha_k}{2}.$$

Proceeding in this fashion we put successively

$$\begin{aligned} |b_0| : \sqrt{(|b_0|^2 + \cdots + |b_k|^2)} &= \cos \alpha_k \geqq 0, \\ |b_1| : \sqrt{(|b_1|^2 + \cdots + |b_k|^2)} &= \cos \alpha_{k-1} \geqq 0, \\ &\vdots \\ |b_{k-1}| : \sqrt{(|b_{k-1}|^2 + |b_k|^2)} &= \cos \alpha_1 \geqq 0, \end{aligned} \tag{5}$$

counting the α's in inverse order. Solving with respect to the b's we have

$$\begin{aligned} |b_0| &= \cos \alpha_k, \\ |b_1| &= \sin \alpha_k \cos \alpha_{k-1}, \\ &\vdots \\ |b_{k-1}| &= \sin \alpha_k \sin \alpha_{k-1} \cdots \cos \alpha_1, \\ \beta = |b_k| &= \sin \alpha_k \sin \alpha_{k-1} \cdots \sin \alpha_1. \end{aligned} \tag{6}$$

From the last equation under (6) it follows furthermore that none of the angles $\alpha_k, \alpha_{k-1}, \cdots, \alpha_1$ can be zero, and (5) completes the information to give

$$0 < \alpha_i \leqq \frac{\pi}{2}.$$

Therefore writing

$$0 < \operatorname{tg} \frac{\alpha_i}{2} = t_i \leqq 1$$

we have the recursion formulae

$$\delta_{i+1} = \delta_i t_i : \{1 + \delta_i^2 t_i^2\}^{\frac{1}{2}}$$

with $\delta_1 = 1$, $\delta_{k+1} = \delta$. Furthermore $\delta_i \leqq 1$; hence

$$\frac{\delta_{i+1}}{\delta_i} \geqq \frac{t_i}{\sqrt{(1 + t_i^2)}} \geqq \frac{t_i}{1 + t_i^2} = \frac{1}{2}\sin\alpha_i$$

and if we finally form the product we obtain

$$\prod_{i=1}^{k} \frac{\delta_{i+1}}{\delta_i} = \frac{\delta_{k+1}}{\delta_1} = \delta \geqq |b_k| \cdot 2^{-k} = 2^{-k}\beta,$$

which together with (7.2) completes the proof of (1).

URBANA, ILLINOIS.

Annals of Mathematics
Vol. 42, No. 2, April, 1941

ON ADDING RELATIONS TO HOMOTOPY GROUPS

By J. H. C. Whitehead

1. Let X be an arcwise connected topological space, and let $\pi_r(X)$ ($r = 1, 2, \cdots$) be the r^{th} homotopy group[1] of X, written with multiplication if $r = 1$ and addition if $r > 1$. Let $f_i(S^{n-1}) \subset X$ ($i = 1, \cdots, k$; $n \geqq 2$) be maps in X of an $(n-1)$-sphere S^{n-1}, let $\mathfrak{E}_i^n$ be a non-singular (open) cell bounded by $f_i(S^{n-1})$, as described below, and let

$$X^* = X + \mathfrak{E}_1^n + \cdots + \mathfrak{E}_k^n \qquad (\mathfrak{E}_i^n \cdot \mathfrak{E}_j^n = 0 \text{ if } i \neq j).$$

In a recent paper[2] I described in algebraical terms the relation between $\pi_{n-1}(X)$ and $\pi_{n-1}(X^*)$, and also the relation between $\pi_n(X)$ and $\pi_n(X^*)$ in case each of $f_i(S^{n-1})$ is homotopic to a point. Here we study the relation between $\pi_n(X)$ and $\pi_n(X^*)$ when the maps $f_i(S^{n-1})$ are arbitrary. There is a considerable difference between the cases $n = 2$ and $n > 2$. In case $n > 2$ the relation between $\pi_n(X)$ and $\pi_n(X^*)$ is expressed in terms of a product $\alpha \cdot \beta \in \pi_{m+n-1}(X)$, where $\alpha \in \pi_m(X)$, $\beta \in \pi_n(X)$. The case $n = 2$ is, in many ways, the more interesting of the two. Among other things a method is found for calculating[3] $\pi_2(K)$ algebraically, where K is any simplicial complex. Of course K. Reidemeister's[4] theory of homology in $\tilde{K}$, the universal covering complex of K, together with a theorem due to W. Hurewicz[5] lead to a theoretical definition of $\pi_2(K)$, which may be stated in purely algebraic terms. But since there is no general algorithm for deciding whether or no given elements $\rho_1, \cdots, \rho_n$ in the group ring, $\mathfrak{R}$, of K, satisfy given equations

$$\sum_j \eta_{\lambda j} \rho_j = 0 \qquad (\eta_{\lambda j} \in \mathfrak{R})$$

this does not lead to a method of calculating $\pi_2(K)$. In fact the problem of calculating the algebraic structure of $\pi_2(K)$ by this method is equivalent to the problem of calculating $\pi_1(K)$ effectively, or of defining $\tilde{K}$ constructively. In order to calculate $\pi_2(K)$ itself one would also need a construction for a deforma-

[1] W. Hurewicz, Kon. Wetensch. Amsterdam, 38 (1935), 112–9; 521–8; 39 (1936), 117–25; 215–24.

[2] Proc. L. M. S., 45 (1939), 243–327, §6. This paper will be referred to as S.S. The argument given in S.S. obviously applies, with minor alterations, when K is any arcwise connected topological space.

[3] I.e., given a constructive definition of K, one can enumerate (constructively) a set of generators and relations for $\pi_2(K)$. Moreover, given a map $f(S^2) \subset K$, with a specified base point, one can express the corresponding element of $\pi_2(K)$ as a product of the generators, and conversely. We shall say that a group G has been calculated effectively when it has been calculated, and when a finite algorithm has been provided for deciding whether or no two given products of the generators represent the same element of G.

[4] See, among papers, K. Reidemeister, Abh. math. Sem. Hamb. 10 (1934), 211–5.

[5] Loc. cit. (Paper II), p. 522.

tion cell, bounded by a given circuit in $\tilde{K}$. It should be said that there is no theoretical obstacle to calculating $\pi_r(K)$, for any $r \geqq 1$, by constructions which are similar to those in a combinatorial definition of[6] $\pi_1(K)$. Thus the value of §6, below, is technical, rather than theoretical, in that it brings new algebraic machinery to bear on the study of $\pi_2(K)$.

We shall always use S^n to denote an oriented n-sphere, and E^n to denote an oriented n-element. The corresponding unoriented spaces will be denoted by $| S^n |$ and $| E^n |$. Let $f(\dot{E}^n) \subset X$ be a given map, where $\dot{E}^n$ is the boundary of E^n and E^n has no point in common with X, and let $\mathfrak{E}^n$ be the interior of E^n. By $X + \mathfrak{E}^n$ we shall mean the space consisting of the topological space X and the topological space $\mathfrak{E}^n$, related by the following condition. If $p_1, p_2, \cdots \subset \mathfrak{E}^n$ is an infinite set of points whose limit points all lie in the closed set $f^{-1}(q) \subset \dot{E}^n$, where q is any point in $f(\dot{E}^n) \subset X$, then the sequence $p_1, p_2, \cdots \subset X + \mathfrak{E}^n$ converges to q. Subject to this condition we describe $\mathfrak{E}^n$ as a non-singular cell bounded by $f(\dot{E}^n)$. When describing a geometrical construction we shall use the term *accidental intersection* to mean one which can be avoided without restricting the conditions of the problem in hand. For example, a point common to $\mathfrak{E}_i^n$ and $\mathfrak{E}_j^n$ $(i \neq j)$, in the above space X^*, would be an accidental intersection. Again, if we introduce a segment $s \subset M^n$, joining two given points p_1, p_2, where M^n is a connected, bounded manifold, then a point, other than p_1 or p_2, which in common to s and $\dot{M}^n$, would be an accidental intersection.

2. In this section we recall some elementary definitions in a convenient form. To define $\pi_n = \pi_n(X)$ we first choose a base point $x_0 \in X$. Then an element $\alpha \in \pi_n$ is given by a map $f(S^n, a) \subset X$, such that $f(a) = x_0$, where a is a specified base point in S^n. In general the same map $f(S^n)$ will represent a different element if another point $a' \in f^{-1}(x_0)$ is chosen as base point in S^n. The element $-\alpha$ is given by $f(-S^n, a)$, where $-S^n$ is S^n with the orientation reversed. Two maps $f_i(S_i^n, a_i)$ $(i = 1, 2)$, where $a_i \in S_i^n$ and $f_i(a_i) = x_0$, represent the same element of π_n if, and only if, $f_1^* = f_2\phi$, where $\phi(S_1^n) = S_2^n$ is a map of degree $+1$ such that $\phi(a_1) = a_2$, and $f_1^*(S_1^n, a_1)$ is homotopic in X, with $f_1(a_1)$ held fixed, to $f_1(S_1^n, a_1)$. Following Hurewicz we may also represent $\alpha \in \pi_n$ by a map $g(E^n) \subset X$ such that $g(\dot{E}^n) = x_0$, in which case we shall always take the map $g(\dot{E}^n) = x_0$ as the base point. Such a map will represent the same element as $f(S^n, a)$ if $g = f\phi$, where $\phi(E^n) = S^n$ is a map of degree $+1$, treating E^n as a cycle mod $\dot{E}^n$, such that $\phi(\dot{E}^n) = a$. Let $\alpha_i \in \pi_n$ $(i = 1, 2; n > 1)$ be given by $f_i(E_i^n) \subset X$, with $f_i(\dot{E}_i^n) = x_0$, and either

1. $| E_1^n | \cdot | E_2^n | = | \dot{E}_1^n | \cdot | \dot{E}_2^n | = | E^{n-1} |$ *and* $\dot{E}_1^n = E^{n-1} + \cdots$, $\dot{E}_2^n = -E^{n-1} + \cdots$, *or*
2. $| E_1^n | \cdot | E_2^n | = | \dot{E}_1^n | = | \dot{E}_2^n |$ *and* $\dot{E}_1^n = -\dot{E}_2^n$.

[6] Since neither can be calculated effectively, except in special cases, the only difference between the logical status of $\pi_1(K)$ and of $\pi_r(K)$ $(r > 1)$ is that $\pi_1(K)$, unlike $\pi_r(K)$, is always given by a finite system of generators and relations if K is finite.

In either case $\alpha_1 + \alpha_2 \in \pi_n$ is the element given by the map $f(E_1^n + E_2^n)$, with a point[7] on $\dot{E}_1^n$ as a base point in the second case, where $f = f_i$ in E_i^n.

Let $S_1^n = E_1^n - E^n$, $S_2^n = E^n - E_2^n$ and $S^n = E_1^n - E_2^n$, where $| E_1^n | \cdot | E_2^n | = | \dot{E}_i^n | = | \dot{E}^n |$. Let f be a given map of $K = | E_1^n | + | E_2^n | + | E^n |$ in X, with $f(a) = x_0$, where $a \in \dot{E}^n$, and let $\alpha_i \in \pi_n$ be the element given by $f(S_i^n, a)$.

LEMMA 1. *The map $f(S^n, a)$ represents the element $\alpha_1 + \alpha_2$.*

This follows at once from the fact that $f(K)$ is homotopic rel. a (i.e. holding $f(a)$ fixed) into a map $f_1(K)$ such that $f_1(E^n) = x_0$.

We recall from S.S., §§10, 11, that π_n $(n > 1)$ is a group with operators. As in S.S., §11, we shall consider the operators to be elements of the group ring $\mathfrak{R} = \mathfrak{R}(\pi_1)$, rather than elements of the ring $\mathfrak{R}_n(X, x_0)$, defined in[8] S.S. §10. If $\alpha \in \pi_n$, $\xi \in \pi_1$ the characteristic property of $\xi\alpha$ is that any of its representative maps can be transformed into a representative of α by a deformation in which the base point describes (positively) a circuit representing ξ. By an invariant sub-group $\sigma_n \subset \pi_n$, we shall mean one such that $\mathfrak{R}\sigma_n = \sigma_n$, that is to say $\rho\alpha \in \sigma_n$ if $\alpha \in \sigma_n$, $\rho \in \mathfrak{R}$. This is a natural generalization of the ordinary definition in case $n = 1$. We shall describe the groups defined on pp. 281 and 283 of S.S., which we shall now denote by $\mathfrak{R}(\alpha_1, \cdots, \alpha_k)$ and $\mathfrak{R}(f_1, \cdots, f_k)$ instead of $r(\alpha_1, \cdots, \alpha_k)$ and $r(f_1, \cdots, f_k)$, as the invariant sub-groups *generated* by $\alpha_1, \cdots, \alpha_k$ and by $f_1, \cdots, f_k$. This definition obviously applies to infinite sets of elements α_i or maps f_i, and in the same way we may define the invariant sub-group generated by a set of elements together with a set of maps.

Let an element $\alpha \in \pi_n$ $(n > 1)$ be given by a map $f(S^n, a) \subset X$ and let $f(a) = f(b) = x_0$. Let $s \subset S^n$ be an oriented segment beginning at b and joining it to a, and let ξ be the element of π_1, with base point x_0, which is represented by the circuit $f(s)$. Then it may be verified that $f(S^n, b)$ represents the element $\xi\alpha$. For this purpose S. Eilenberg's definition[9] of $\xi\alpha$ is particularly convenient.

3. Let $\alpha \in \pi_m = \pi_m(X)$ and $\beta \in \pi_n = \pi_n(X)$ $(m, n \geqq 1)$ be given elements represented by maps $f_0(E^m) \subset X$ and $g_0(E^n) \subset X$, such that $f_0(\dot{E}^m) = g_0(\dot{E}^n) = x_0$, where x_0 is to be taken as the base point for all the homotopy groups $\pi_r(X)$. We denote by $f_0 \cdot g_0$ the map

$$F_0(E^m \times E^n)^{\cdot} = F_0(\dot{E}^m \times E^n + (-1)^m E^m \times \dot{E}^n) \subset X,$$

given by

$$\begin{aligned} F_0(p \times q) &= g_0(q) \quad \text{if} \quad p \in \dot{E}^m, q \in E^n \\ &= f_0(p) \quad \text{if} \quad p \in E^m, q \in \dot{E}^n. \end{aligned}$$

---

[7] Since $n > 1$, $\dot{E}_1^n$ is connected and it is immaterial which point on E_1^n is taken as base point (see the concluding remark in this section).

[8] The distinction is that, if $\rho_1\alpha = \rho_2\alpha$ for each $\alpha \in \pi_n$, then ρ_1 and ρ_2 are identical, by definition, if regarded as elements of $\mathfrak{R}_n(X, x_0)$, but they may be different elements of $\mathfrak{R}$.

[9] Fund. Math., 32 (1939), 167–75. Eilenberg defines $\xi\alpha$ in terms of the universal covering space of X.

We take a point $a \times b \in \dot{E}^m \times \dot{E}^n$ as base point on $(E^m \times E^n)^{\cdot}$, where a is an arbitrary point on $\dot{E}^m$ if $m > 1$, and the first point of $\dot{E}^m$ if $m = 1$, and similarly for $b \in \dot{E}^n$. Let f_1 and g_1 be the images of f_0 and g_0 in homotopic deformations f_t and g_t $(0 \leqq t \leqq 1)$, such that $f_t(\dot{E}^m) = g_t(\dot{E}^n) = x_0$ for each $t \in \langle 0, 1\rangle$. Then $f_1 \cdot g_1 = F_1$ is the image of F_0 in the deformation F_t, given by

$$F_t(p \times q) = g_t(q) \quad \text{if} \quad p \in \dot{E}^m,\ q \in E^n$$

$$= f_t(p) \quad \text{if} \quad p \in E^m,\ q \in \dot{E}^n,$$

throughout which $F_t(a \times b) = x_0$. Therefore the element of $\pi_{m+n-1}(X)$ determined by the map $f_0 \cdot g_0$ depends only on the elements $\alpha \in \pi_m$, $\beta \in \pi_n$. We shall denote it by $\alpha \cdot \beta$. If $m = n = 1$ it is clear that

$$\xi \cdot \eta = \xi\eta\xi^{-1}\eta^{-1} \qquad (\xi, \eta \subset \pi_1), \tag{3.1}$$

and if $m = 1$, $n > 1$, that

$$\xi \cdot \beta = (\xi - 1)\beta. \tag{3.2}$$

If $m + n \geqq 2$ we have

$$\beta \cdot \alpha = (-1)^{mn} \alpha \cdot \beta \tag{3.3}$$

since $E^n \times E^m = (-1)^{mn} E^m \times E^n$.

THEOREM 1. *If $n > 1$ the transformation $\beta \to \alpha \cdot \beta$ is a homomorphism of π_n in π_{m+n-1} for each $\alpha \in \pi_m$.*

Let $\beta = \beta_1 + \beta_2$, where $\beta_i \in \pi_n$ $(i = 1, 2)$. Let β_i be represented by a map $g_i(E_i^n) \subset X$, where

$$|E_1^n| \cdot |E_2^n| = |\dot{E}_1^n| \cdot |\dot{E}_2^n| = |E^{n-1}|,$$

$$\dot{E}_1^n = E^{n-1} + \cdots, \qquad \dot{E}_2^n = -E^{n-1} + \cdots,$$

and $g_i(\dot{E}_i^n) = x_0$. Then β is given by $g(E^n)$, where $E^n = E_1^n + E_2^n$ and $g = g_i$ in E_i^n, and since $n > 1$ we may take a point in $\dot{E}^m \times E^{n-1}$ as the base point on $(E^m \times E^n)^{\cdot}$. Then the two $(m + n - 1)$-spheres

$$S_i^{m+n-1} = \dot{E}^m \times E_i^n + (-1)^m E^m \times \dot{E}_i^n$$

meet in the $(m + n - 1)$-element $E^m \times E^{n-1}$, and

$$S_1^{m+n-1} + S_2^{m+n-1} = \dot{E}^m \times E^n + (-1)^m E^m \times \dot{E}^n.$$

Therefore $\alpha \cdot \beta_1 + \alpha \cdot \beta_2 = \alpha \cdot (\beta_1 + \beta_2)$, by lemma 1, and the theorem is established.

If $m > 1$, $n > 1$ it follows from theorem 1 and (3.3) that the function $\alpha \cdot \beta$ determines a group multiplication.[10] However it may happen that $\alpha \cdot \beta = 0$ even though $\alpha \neq 0$, $\beta \neq 0$, as it does when $m = 1$ and X is n-simple in the sense of Eilenberg.[11]

---

[10] Cf. P. Alexandroff and H. Hopf, *Topologie*, Berlin (1935), 589–90.

[11] Loc. cit.

The homomorphism $\beta \to \alpha \cdot \beta$ ($n > 1$) is not, in general, an operator homomorphism with respect to the operators in $\mathfrak{R}$. In fact if $m, n > 1$ it may be verified that

$$\xi(\alpha \cdot \beta) = \xi\alpha \cdot \xi\beta \qquad (\xi \in \pi_1).$$

If $m = 1$, $n > 1$ we have

$$\begin{aligned}\xi(\eta \cdot \beta) &= \xi(\eta - 1)\beta \\ &= (\xi\eta\xi^{-1} - 1)\xi\beta \\ &= \xi\eta\xi^{-1} \cdot \xi\beta\end{aligned}$$

for any $\eta \in \pi_1$, and if $m = n = 1$ the relation

$$\xi(\eta \cdot \eta')\xi^{-1} = \xi\eta\xi^{-1} \cdot \xi\eta'\xi^{-1} \qquad (\eta, \eta' \subset \pi_1)$$

follows from (3.1).

The product $\alpha \cdot \beta$ is one among many similar ways in which elements $\alpha \in \pi_m$ and $\beta \in \pi_n$ may be combined to form an element in π_{m+n-1}. For let α and β be given by $f(E^m) \subset X$ and $g(E^n) \subset X$, with $f(\dot{E}^m) = g(\dot{E}^n) = x_0$, where E^m and E^n are now the interior regions and boundaries of Euclidean metric spheres S^{m-1} and S^{n-1} in Euclidean spaces R^m and R^n. Let $p \to r_p$ ($p \in S^{m-1}$, $r_p \in G_n$) be a continuous map of S^{m-1} in G_n, the group of rotations in R^n about the center of S^{n-1}, and let $q \to r_q$ ($q \in S^{n-1}$, $r_q \in G_m$) be similarly defined. Then $\alpha \cdot \beta$ is a special case of the elements given by maps of the form $F(E^m \times E^n) \subset X$, where

$$\begin{aligned}F(p \times q) &= g\{r_p(q)\} \quad \text{if} \quad p \in \dot{E}^m,\ q \in E^n \\ &= f\{r_q(p)\} \quad \text{if} \quad p \in E^m,\ q \in \dot{E}^n.\end{aligned}$$

Let $X = S^n$ ($n \geqq 2$), let $m = 2$ let $\alpha = 0$, even when $n = 2$, and let $g(E^n) \subset S^n$, with $g(\dot{E}^n) = x_0$, be of degree 1. Then if $n = 2$ it may be verified, with the help of H. Hopf's invariant,[12] that the map $p \to r_p$ ($p \in S^1$) determines an isomorphism between $\pi_1(G_n)$ and $\pi_{n+1}(S^n)$. The fact that $\pi_1(G_n)$ and $\pi_{n+1}(S^n)$ are both cyclic infinite if $n = 2$ and of order two[13] if $n > 2$ suggests that the same may be true if $n > 2$.

4. Let S_1^m and S_2^n ($1 < m \leqq n$) be m- and n-spheres with a single common point b, and let $r = m + n - 1 > 1$. Let $\pi_{r1} = \pi_r(S_1^m)$, $\pi_{r2} = \pi_r(S_2^n)$ and let $\pi_m \cdot \pi_n$ be the sub-group of $\pi_r = \pi_r(S_1^m + S_2^n)$ which is generated by all elements of the form $\alpha \cdot \beta$, where $\alpha \in \pi_m(S_1^m)$, $\beta \in \pi_n(S_2^n)$ and b is the base point of π_{ri}, $\pi_m \cdot \pi_n$ and π_r.

THEOREM 2. *The group π_r is the direct sum $\pi_r(S_1^m) + \pi_r(S_2^n) + \pi_m \cdot \pi_n$. The group $\pi_m \cdot \pi_n$ is cyclic infinite.*

[12] Math. Ann., 104 (1931), 637-65.

[13] L. Pontrjagin, Oslo Congress (1936): H. Freudenthal, Compositio Math., 5 (1938), 299-314.

Let[14] $f(S^r) \subset S_1^m + S_2^n$ be a map representing a given element $\alpha^* \epsilon \pi_r$. We take S_1^m, S_2^n and S^r to be recti-linear sub-divisions of the boundaries of recti-linear simplexes, and we assume that the map f is simplicial. Let x be an inner point of an n-simplex in S_1^n. Then $f^{-1}(x)$ is a polyhedron whose cells may be oriented to form an $(m-1)$-cycle $Z^{m-1} \subset S^r$. Let $Z^m \subset S^r$ be a chain bounded by Z^{m-1} and let $\gamma_1(f)$ be the degree with which $f(Z^m)$ covers S_1^m. It follows from his original argument,[15] with trivial modifications, that $\gamma_1(f)$ is a 'Hopf invariant' of the element α^*, and it may therefore be written $\gamma_1(\alpha^*)$. Clearly $\alpha^* \to \gamma_1(\alpha^*)$ is a homomorphism of π_r in the additive group of integers.

We now prove the last part of the theorem. The groups $\pi_m(S_1^m)$ and $\pi_n(S_2^n)$ are cyclic infinite, generated by elements α and β, say. Therefore any element in $\pi_m \cdot \pi_n$ is of the form $k\alpha \cdot l\beta = kl(\alpha \cdot \beta)$, by theorem 1, where k and l are integers. Therefore $\pi_m \cdot \pi_n$ is a cyclic group generated by $\alpha \cdot \beta$. Obviously

$$\gamma_1(k\alpha \cdot l\beta) = \pm kl,$$

whence $k\alpha \cdot l\beta = 0$ implies $k = 0$ or $l = 0$. Taking $l = 1$ it follows that $k(\alpha \cdot \beta) \neq 0$ if $k \neq 0$, or that $\pi_m \cdot \pi_n$ is infinite.

We have now to prove that $\pi_r = \pi_{r1} + \pi_{r2} + \pi_m \cdot \pi_n$. Let $h_1(S_1^m + S_2^n) = S_1^m$ be the map given by

$$\begin{aligned} h_1(p) &= p \quad \text{if} \quad p \epsilon S_1^m \\ &= b \quad \text{if} \quad p \epsilon S_2^n. \end{aligned}$$

Let $\phi_1(\pi_r) = \pi_{r1}$ be the homomorphism of π_r in which each element given by a map $f(S^r) \subset S_1^m + S_2^n$ is transformed into the element given by the map $h_1 f(S^r)$. Clearly $\phi_1(\pi_{r2}) = 0$ and $\phi_1(\alpha) = \alpha$ if $\alpha \epsilon \pi_{r1}$. Since $\alpha \cdot 0 = 0$, according to theorem 1, we have $\phi_1(\pi_m \cdot \pi_n) = 0$. If $\alpha_1 + \alpha_2 + \alpha \cdot \beta = 0$, where $\alpha_i \epsilon \pi_{ri}$, $\alpha \epsilon \pi_m(S_1^m)$, $\beta \epsilon \pi_n(S_2^n)$, it follows that

$$\alpha_1 = \phi_1(\alpha_1 + \alpha_2 + \alpha \cdot \beta) = 0.$$

Similarly $\alpha_2 = 0$ and hence $\alpha \cdot \beta = 0$. Therefore the group consisting of all elements of the form $\alpha_1 + \alpha_2 + \alpha \cdot \beta$ is the direct sum $\pi_{r1} + \pi_{r2} + \pi_m \cdot \pi_n$, and it remains to show that each element in π_r is of this form.

Let $\gamma_1(\alpha^*) = k$, where α^* is a given element in π_r, and, replacing α by $-\alpha$ if necessary, let $\gamma_1(\alpha \cdot \beta) = 1$, where $\alpha \cdot \beta$ generates $\pi_m \cdot \pi_n$. Then $\gamma_1(\alpha_0) = 0$, where $\alpha_0 = \alpha^* - k\alpha \cdot \beta$. Let $f(S^r) \subset S_1^m + S_2^n$ be a map representing α_0 and, as before, let $\dot{Z}^m = Z^{m-1}$, where $| Z^{m-1} | = f^{-1}(x)$. If $n > m$ it follows from a fundamental theorem due to Hopf[16] that $f(Z^m)$ can be deformed into a point, namely $f(Z^{m-1}) = x$, with $f(Z^{m-1})$ held fixed. Therefore an argument used by H. Freudenthal[17] shows that α_0 may be represented by a simplicial map f_1 such that $f_1^{-1}(x)$ is a single point p. Let $E^n \subset S_2^n$ and $E^r \subset S^r$ be the simplicial

[14] Since $\pi_1(S_1^m + S_2^n) = 1$ we need not specify which point in $f^{-1}(b)$ is to be taken as the base point in S^r.

[15] H. Hopf, loc. cit.: also Fund. Math., 25 (1935), 427-40.

[16] Comment. math. helv., 5 (1933), 39-54.

[17] Loc. cit., pp. 309-11.

neighborhoods of x and p, and let $\psi(S_2^n) = S_2^n$ be a map of degree unity such that $\psi(S_2^n - E^n) = b$, $\psi(E^n) = S_2^n$. We extend ψ throughout $S_1^m + S_2^n$ by writing $\psi(\eta) = y$ if $y \in S_1^m$. Clearly $\psi f_1(S^r)$ is homotopic to f_1, with the base-point, in $f_1^{-1}(b)$, held fixed. Therefore α_0 is also represented by the map ψf_1. But $\psi f_1(E^r) \subset S_2^n$, $\psi f_1(E_1^r) \subset S_1^m$ and $\psi f_1(S^{r-1}) = b$, where E_1^r is the closure of $S^r - E^r$ and $| S^{r-1} | = | \dot{E}^r | = | \dot{E}_1^r |$. Therefore $\alpha_0 = \alpha_1 + \alpha_2$ and $\alpha^* = \alpha_1 + \alpha_2 + k\alpha\cdot\beta$, where $\alpha_1 \in \pi_{r1}$ is represented by $\psi f_1(E_1^r) \subset S_1^m$, and $\alpha_2 \in \pi_{r2}$ by $\psi f_1(E^r) \subset S_2^n$. Therefore the theorem is established in case $m < n$.

Finally let $m = n$. With the same notation as before, let $\alpha^* = \alpha_0 + k\alpha\cdot\beta$, where α^* is a given element in π_r and $\gamma_1(\alpha_0) = 0$, and let $\gamma_2(\alpha_0)$ be the degree with which $f(Z^m)$ covers S_2^n. Clearly $\gamma_1\{\phi_2(\alpha_0)\} = 0$, $\gamma_2\{\phi_2(\alpha_0)\} = \gamma_2(\alpha_0)$, where ϕ_2 is the homomorphism induced by the map h_2, given by $h_2(S_1^m) = b$, $h_2(p) = p$ if $p \in S_2^n$. Therefore $\gamma_1(\alpha_0^*) = \gamma_2(\alpha_0^*) = 0$, where $\alpha_0^* = \alpha_0 - \phi_2(\alpha_0)$, and it follows as when $n > m$ that $\alpha_0^* \in \pi_{r1} + \pi_{r2}$. But $\phi_2(\alpha_0) \in \pi_{r2}$ and $\alpha^* - \alpha_0 \in \pi_m\cdot\pi_n$. Therefore $\alpha^* = \alpha_0^* + (\alpha_0 - \alpha_0^*) + (\alpha^* - \alpha_0) \in \pi_{r1} + \pi_{r2} + \pi_m\cdot\pi_n$, and the proof is complete.

5. Let $f_i(S_i^{n-1}) \subset X_0$ $(i = 1, \cdots, k)$ be given maps of oriented $(n-1)$-spheres $S_1^{n-1}, \cdots, S_k^{n-1}$ in an arcwise connected space X_0. Let

$$X^* = X_0 + \mathfrak{E}_1^n + \cdots + \mathfrak{E}_k^n, \tag{5.1}$$

where $\mathfrak{E}_i^n$ is a non-singular cell bounded by $f_i(S_i^{n-1})$ and there are no accidental intersections. Let a_i^n be an open n-simplex, oriented in agreement with $\mathfrak{E}_i^n$, such that $A_i^n \subset \mathfrak{E}_i^n$, where $A_i^n = \bar{a}_i^n$, the closure of a_i^n. Then X_0 is a retract by deformation of

$$X = X_0 + \sum_{i=1}^k (\mathfrak{E}_i^n - a_i^n),$$

and it follows that the homotopy groups $\pi_r(X_0)$ and $\pi_r(X)$ are the same, likewise the relations between $\pi_r(X_0)$ and $\pi_r(X^*)$ and between $\pi_r(X)$ and $\pi_r(X^*)$. Also the identical map of $\dot{A}_i^n$ on itself is homotopic in X to $f_i(S_i^{n-1})$. Therefore we may replace X_0 by X and $f_i(S_i^{n-1})$ by the identical map of $\dot{A}_i^n$ on itself. Since $\mathfrak{E}_i^n$ may be triangulated and $A_i^n \subset \mathfrak{E}_i^n$, we may assume, after a suitable deformation, that any map $f(K) \subset X^*$, of a simplicial complex K, is simplicial in $f^{-1}(A_1^n + \cdots + A_k^n)$. From here to the last paragraph in this section we take $n > 2$.

Let $x_0 \in X$ be a base point for each $\pi_r = \pi_r(X)$ and let α_i be the element in π_{n-1} which is given by[18] $t_i + \dot{A}_i^n$, where t_i is an oriented segment starting at x_0 and joining it to a point $x_i \in \dot{A}_i^n$. Let $\mathfrak{M}$ be a modulus with independent basis elements $e_1, \cdots, e_k$ and coefficients in $\mathfrak{R} = \mathfrak{R}(\pi_1)$. Then the transformation ϕ, given by

$$\phi(\rho_1 e_1 + \cdots + \rho_k e_k) = \rho_1\alpha_1 + \cdots + \rho_k\alpha_k \qquad (\rho_i \in \mathfrak{R}) \tag{5.2}$$

[18] Cf. S.S., p. 279.

is a homomorphism of $\mathfrak{M}$ on $\mathfrak{R}(\alpha_1, \cdots, \alpha_k)$. Clearly $\mathfrak{R}\mathfrak{M}_0 = \mathfrak{M}_0$, where $\mathfrak{M}_0 = \phi^{-1}(0)$ is the kernel of ϕ.

Let $f(S^n, p_0) \subset X^*$, with $f(p_0) = x_0$, be given. After a suitable deformation, we assume that $f^{-1}(A_i^n)$ is a set of oriented n-simplexes $\epsilon_{i1}A_{i1}^n, \cdots, \epsilon_{iq_i}A_{iq_i}^n$, where $\epsilon_{i\lambda} = \pm 1$ $(i = 1, \cdots, k)$, $A_{i\lambda}^n$ takes its orientation from S^n, and $f(A_{i\lambda}^n) = \epsilon_{i\lambda}A_i^n$, the map f being linear in $A_{i\lambda}^n$. After a suitable sub-division of S^n and a further deformation[19] of f we assume that no two of the simplexes $A_{i\lambda}^n, A_{j\mu}^n$ have a common point. Let $s_{i\lambda} \subset S^n$ be an oriented, polygonal segment, starting at p_0 and joining it, without accidental intersections, to the point $p_{i\lambda} \in \dot{A}_{i\lambda}^n$, such that $f(p_{i\lambda}) = x_i$, and let $t_{i\lambda} = f(s_{i\lambda})$. Then $f(s_{i\lambda} + \dot{A}_{i\lambda}^n) = t_{i\lambda} + \epsilon_{i\lambda}\dot{A}_i^n$ and is homotopic to $(t_{i\lambda} - t_i) + (t_i + \epsilon_{i\lambda}\dot{A}_i^n)$. Therefore the corresponding element in π_{n-1} is $\epsilon_{i\lambda}\xi_{i\lambda}\alpha_i$, where $\xi_{i\lambda} \in \pi_1$ is given by the circuit $t_{i\lambda} - t_i$. Then

$$\sum_{i,\lambda} \epsilon_{i\lambda}\xi_{i\lambda}\alpha_i = 0 \tag{5.3}$$

since $\sum_{i,\lambda}(t_{i\lambda} + \epsilon_{i\lambda}\dot{A}_i^n) = f\{\sum_{i,\lambda}(s_{i\lambda} + \dot{A}_{i\lambda}^n)\}$, and the singular sphere $\sum_{i,\lambda}(s_{i\lambda} + \dot{A}_{i\lambda}^n)$ bounds the cell $S^n - \sum_{i,\lambda}(s_{i\lambda} + A_{i\lambda}^n)$. Therefore

$$\psi(f) = \sum_{i,\lambda} \epsilon_{i\lambda}\xi_{i\lambda}e_i \in \mathfrak{M}_0. \tag{5.4}$$

Thus to the map f corresponds an element $\psi(f) \in \mathfrak{M}_0$. If $f(S^n) \subset X$, the set of simplexes $f^{-1}(A_i^n)$ being empty, we set $\psi(f) = 0$.

LEMMA 2. *The element $\psi(f) = \psi(\alpha^*)$ depends only on the element $\alpha^* \in \pi_n^* = \pi_n(X^*)$, which is given by f.*

In the definition of $\psi(f)$, and in proving the lemma, it is obvious that a map of the form $f(S^n) \subset X^*$ may at any stage be replaced by one of the form $f(E^n) \subset X^*$, with $f(\dot{E}^n) = x_0$. Therefore the lemma will follow if we can show that $\psi(f_1) = \psi(f_2)$, where $f_i(E_i^n)$ $(i = 1, 2)$ are two given maps representing the same element $\alpha^* \in \pi_n^*$, both of which are simplicial in $f^{-1}(A_1^n + \cdots + A_k^n)$. We assume as we obviously may, that E_1^n and E_2^n are two hemispheres of an n-sphere $S^n = E_1^n - E_2^n$, where $\dot{E}_1^n = \dot{E}_2^n$, with regard to orientation. Then the map $f(S^n, p_0) \subset X^*$, where $p_0 \in \dot{E}_1^n$ and $f = f_i$ in E_i^n, represents the element $\alpha^* - \alpha^* = 0$. Clearly $\psi(f) = \psi(f_1) - \psi(f_2)$. So we have to show that $\psi(f) = 0$. Let $S^n = \dot{E}^{n+1}$. Then the map f can be extended throughout E^{n+1}, and we assume that $f(E^{n+1})$ is simplicial[20] in $f^{-1}(A_1^n + \cdots + A_k^n)$.

---

[19] Cf. S.S., p. 282.

[20] For it is given that f is simplicial in $f^{-1}(A_1^n + \cdots + A_k^n) \subset S^n$, and, among the processes of sub-division and canonical displacement of vertices by which $f(E^{n+1})$ is made simplicial in $f^{-1}(A_1^n + \cdots + A_k^n)$, there is obviously one which leaves $\psi(f)$ unaltered. Alternatively we can appeal to the following theorem, which is easy to prove, though I have not seen it in print. Let K and L be simplicial complexes and let $f_0(K) \subset L$ be a map which is simplicial in some sub-complex $K^* \subset K$. Then there is a sub-division K_1, of K, which leaves K^* untouched, and a deformation $f_t(K) \subset L$ $(0 \leqq t \leqq 1)$, such that $f_t = f_0$ in K^* and f_1 is simplicial with respect to K_1.

If $f(S^n) \subset X$ it is trivial that $\psi(f) = 0$. Otherwise let $f(A^n_{i\lambda}) = \epsilon_{i\lambda}A^n_i$, where $A^n_{i\lambda} \subset \dot{E}^{n+1}$, let y_i be an inner point of A^n_i, and let s be the segment in $f^{-1}(y_i)$ which, starting at $A^n_{i\lambda}$, terminates at some $A^n_{i\mu} \subset \dot{E}^{n+1}$. Let C^{n+1} be the chain of oriented $(n + 1)$-simplexes in $f^{-1}(A^n_i)$ which contain points of s. Then $\dot{E}^{n+1} = \dot{C}^{n+1} + \cdots$ and also $\dot{E}^{n+1} = A^n_{i\lambda} + A^n_{i\mu} + \cdots$. Therefore $\dot{C}^{n+1} = A^n_{i\lambda} + A^n_{i\mu} + \cdots$. Moreover $f(\dot{C}^{n+1} - A^n_{i\lambda} - A^n_{i\mu}) \subset \dot{A}^n_i$. Since $f(C^{n+1}) = A^n_i$, whence $f(\dot{C}^{n+1})$ is algebraically zero, we have $f(A^n_{i\lambda}) = -f(A^n_{i\mu})$, or $\epsilon_{i\lambda} = -\epsilon_{i\mu}$. Also $p_{i\lambda} \epsilon \dot{A}^n_{i\lambda}$ is joined to $p_{i\mu} \epsilon \dot{A}^n_{i\mu}$ by a segment $s^* \subset \dot{C}^{n+1}$, such that $f(s^*) = x_i$. Since the circuit $s_{i\lambda} + s^* - s_{i\mu}$ bounds a cell in E^{n+1}, and since $f(s^*) = x_i$, the circuit $f(s_{i\lambda}) - f(s_{i\mu})$ bounds a cell in X^*, and hence in X, since $n > 2$. Therefore $\xi_{i\lambda} = \xi_{i\mu}$. On repeating this argument it follows that, for each $i = 1, \cdots, k$, the simplexes $A^n_{i\lambda} \subset \dot{E}^{n+1}$ occur in pairs $A^n_{i\lambda}, A^n_{i\mu_\lambda}$, such that $\epsilon_{i\lambda} = -\epsilon_{i\mu_\lambda}$ and $\xi_{i\lambda} = \xi_{i\mu_\lambda}$. Therefore the expression (5.4) can be reduced to zero by cancelling terms of the form $\epsilon_{i\lambda}(\xi_{i\lambda} + \xi_{i\mu_\lambda})e_i$, where $\xi_{i\mu_\lambda} = \xi_{i\lambda}$. Therefore $\psi(f) = 0$, and the lemma is established.

It follows from an argument in the proof of lemma 2 that $\alpha^* \to \psi(\alpha^*)$ is a homomorphism of π^*_n in $\mathfrak{M}_0$.

LEMMA 3. *ψ is a homomorphism on $\mathfrak{M}_0$. It is an operator homomorphism, meaning that $\psi(\rho\alpha^*) = \rho\psi(\alpha^*)$ for any $\rho \epsilon \mathfrak{R}$. Its kernel is the sub-group $\pi^0_n \subset \pi^*_n$, which consists of elements with representative maps in X.*

Let α, given by (5.3), be a given element in $\mathfrak{R}(\alpha_1, \cdots, \alpha_k)$ and let $s_{i\lambda} + A^n_{i\lambda} \subset S^n$ mean the same as before. Then α is represented by a map $f(\Sigma) \subset X$, where

$$\Sigma = \sum_{i,\lambda} (s_{i\lambda} + \dot{A}^n_{i\lambda}),$$

such that $f(\dot{A}^n_{i\lambda}) = \epsilon_{i\lambda}\dot{A}^n_i$, $f(p_0) = x_0$, $f(p_{i\lambda}) = x_i$, and the circuit $f(s_{i\lambda}) - t_i$ represents the element $\xi_{i\lambda} \epsilon \pi_1$. The singular sphere Σ may be regarded as the boundary of the cell $S^n - \sum_{i,\lambda} (s_{i\lambda} + A^n_{i\lambda})$. Let e, given by (5.4), be an arbitrary element of $\mathfrak{M}_0$. Then $\alpha = 0$ and the map $f(\Sigma)$ can therefore be extended to a map $f(S^n - \sum_{i,\lambda} a^n_{i\lambda}) \subset X$, where $a^n_{i\lambda}$ is the interior of $A^n_{i\lambda}$, and hence to a map $f(S^n) \subset X^*$. Then $\psi(\alpha^*) = e$, where $\alpha^* \epsilon \pi^*_n$ is the element given by $f(S^n, p_0)$, and it follows that ψ is a homomorphism of π^*_n on $\mathfrak{M}_0$.

Let $\alpha^* \epsilon \pi^*_n$ be represented by a map $f(E^n) \subset X^*$, with $f(\dot{E}^n) = x_0$, and let f be simplicial in $f^{-1}(A^n_1 + \cdots + A^n_k)$. Let $A^n_{i\lambda}$ mean the same as before, but now let $s_{i\lambda}$ be a segment which, without accidental intersections, joins a point $p'_{i\lambda} \epsilon \dot{E}^n$ to $p_{i\lambda}$, where the points $p'_{11}, p'_{12}, \cdots$ are distinct. Let E^n_0 be an n-element such that $| E^n | \cdot | E^n_0 | = | \dot{E}^n | = | \dot{E}^n_0 |$, let p'_0 be an inner point of E^n_0 and let $s'_{i\lambda} \subset E^n_0$ be an oriented segment which starts at p'_0 and joins it to $p'_{i\lambda}$. Let ξ be any element in π_1, and let ξ be given by a map $g(b'b) \subset X$, with $g(b') = g(b) = x_0$, where $b'b$ is a simple segment. Let $f_0(E^n_0) = b'b$ be a map such that $f_0(p'_0) = b'$, $f_0(\dot{E}^n) = b$. Then $\xi\alpha^* \epsilon \pi^*_n$ is given by the map $f^*(S^n, p'_0) \subset X^*$, where $S^n = E^n + E^n_0$, $f^* = f$ in E^n and $f^* = gf_0$ in E^n_0. The circuit $f^*(s'_{i\lambda} + s_{i\lambda}) - t_i = f^*(s'_{i\lambda}) + f(s_{i\lambda}) - t_i$ represents the element $\xi\xi_{i\lambda} \epsilon \pi_1$.

Therefore $\psi(\xi\alpha^*) = \xi\psi(\alpha^*)$ and $\psi(\rho\alpha^*) = \rho\psi(\alpha^*)$, for any $\rho \in \mathfrak{R}$, since ψ is a homomorphism.

Clearly $\psi(\pi_n^0) = 0$. Conversely, let $\alpha^* \in \psi^{-1}(0)$ be given by a map $f(S^n, p_0) \subset X^*$, which is simplicial in $f^{-1}(A_1^n + \cdots + A_k^n)$. Then, since $e_1, \cdots, e_k$ are linearly independent, the expression (5.4) can be reduced to zero by cancelling pairs of terms of the form $(\xi_{i\lambda} - \xi_{i\mu})e_i$, where $\xi_{i\lambda} = \xi_{i\mu}$. If $\xi_{i\lambda} = \xi_{i\mu}$ the circuit $f(-s_{i\lambda} + s_{i\mu}) = -f(s_{i\lambda}) + f(s_{i\mu})$ is homotopic to a point in X. Therefore the circuit $f(s)$ can be deformed into a point, where s is any segment joining $p_{i\lambda}$ to $p_{i\mu}$, without accidental intersections. Let $E_1^n, E_2^n \subset S^n$ be n-elements such that

$$A_{i\lambda}^n + s + A_{i\mu}^n \subset \mathfrak{E}_1^n, \qquad E_1^n \subset \mathfrak{E}_2^n$$

and E_2^n does not meet p_0 or $A_{j\nu}^n$ $(j, \nu \neq i, \lambda$ or $i, \mu)$, where $\mathfrak{E}_h^n$ is the interior of E_h^n $(h = 1, 2)$. Then it follows from a standard argument[21] that $f(S^n)$ is homotopic, rel. $(S^n - \mathfrak{E}_2^n)$, to a map $f_1(S^n)$, such that $f_1(E_1^n) = x_i$, $f_1(E_2^n - \mathfrak{E}_1^n) \subset X$. Therefore the lemma follows from induction on the number of simplexes in $f^{-1}(A_1^n + \cdots + A_k^n)$.

Let α be any element of $\pi_n = \pi_n(x)$, given by a map $f(S^n) \subset X$, and let $\psi(\alpha) \in \pi_n^0$ be the element of π_n^* which is given by the same map. Then $\alpha \to \psi(\alpha)$ is obviously a homomorphism of π_n on π_n^0. Let σ_n be the invariant sub-group of π_n which is generated by maps of the form $f(S^n) \subset \dot{A}_i^n$ $(i = 1, \cdots, k)$ together with all elements of the form $\alpha_i \cdot \beta$, where $\alpha_1, \cdots, \alpha_k$ mean the same as before and $\beta \in \pi_2(X)$.

LEMMA 4. *σ_n is the kernel of the homomorphism $\psi(\pi_n) = \pi_n^0$.*

Clearly $\psi(\sigma_n) = 0$, and we have to show that, given a map $f(E^{n+1}) \subset X^*$ with $f(\dot{E}^{n+1}) \subset X$, then the class of elements $\mathfrak{R}\{f(\dot{E}^{n+1})\} \subset \pi_n$ is contained in σ_n. Let $y_i \in a_i^n$ and let the given map $f(E^{n+1})$ be simplicial in $f^{-1}(A_1^n + \cdots + A_k^n)$. Since $f(\dot{E}^{n+1}) \subset X$ it follows that $f^{-1}(y_1 + \cdots + y_k)$ is a set of simple, non-intersecting, polygonal circuits inside E^{n+1}. I say that the circuits in $f^{-1}(y_1 + \cdots + y_k)$ bound a set of non-singular, non-intersecting 2-elements. This is certainly the case if $n + 1 > 4$ since singularities and intersections between 2-cells can then be eliminated by slight deformations. If $n + 1 = 4$ $(n + 1 \geqq 4)$ since $n > 2)$ let s be any circuit in $f^{-1}(y_i)$. Then $s = \dot{E}^2$, where $E^2 \subset E^{n+1}$ is a non-singular, polyhedral 2-element, which does not meet $f^{-1}(y_1 + \cdots + y_k) - s$. For a suitable sub-division of E^{n+1} may be represented as a rectilinear sub-division of a rectilinear $(n + 1)$-simplex, and we may take E^2 to be a star whose center is in general position relative to $f^{-1}(y_1 + \cdots + y_k)$. Let E_0^{n+1} be a regular neighborhood[22] of E^2 which does not meet $\dot{E}^{n+1}$ or any of the other circuits in $f^{-1}(y_1 + \cdots + y_k)$. Then $\dot{E}_0^{n+1}$ may be joined to $\dot{E}^{n+1}$ by an $(n+1)$-element $E_1^{n+1} \subset E^{n+1} - f^{-1}(y_1 + \cdots + y_k)$, such that $| E_1^{n+1} | \cdot | E_0^{n+1} | = | \dot{E}_1^{n+1} | \cdot | \dot{E}_0^{n+1} | = | E_0^n |$ and $| E_1^{n+1} | \cdot | E^{n+1} | = | \dot{E}_1^{n+1} | \cdot | \dot{E}^{n+1} | = | E^n |$. The

---

[21] See, for example, Alexandroff and Hopf (loc. cit.), p. 503; or S. Lefschetz, Fund. Math., 27 (1936), 94–115 (pp. 99–100).

[22] S.S., p. 293.

closure of $|E^{n+1}| - |E_1^{n+1}|$ meets E_1^{n+1} in the closure of $|\dot{E}^{n+1}| - |E^n|$, which is an n-element with no internal simplex in $\dot{E}^{n+1}$. Therefore the closure of $|E^{n+1}| - |E_1^{n+1}|$ is an $(n + 1)$-element. Similarly the closure of $|E^{n+1}| - |E_1^{n+1}| - |E_0^{n+1}|$ is an $(n + 1)$-element and the assertion follows from induction on the number of circuits in $f^{-1}(y_1 + \cdots + y_k)$. In the preceding argument, which also applies when $n + 1 > 4$, we cannot be certain that $f(\dot{E}_0^{n+1}) \subset X$, only that $f(\dot{E}_0^{n+1}) \subset X^* - (y_1 + \cdots + y_k)$. But we may replace A_i^n by a simplex $B_i^n \subset A_i^n$, such that $y_i \in B_i^n - \dot{B}_i^n$ and $f(\dot{E}_0^{n+1}) \subset X^* - (B_1^n + \cdots + B_k^n)$. This does not alter π_n, σ_n or π_n^0, or the relations between them. Therefore the lemma will follow from induction on the number of circuits in $f^{-1}(y_1 + \cdots + y_k)$ if we can show that $\Re\{f(\dot{E}_0^{n+1})\} \subset \sigma_n$, where A_i^n is replaced by B_i^n in the definition of σ_n. However, rather than this, we shall simplify the notation by starting again with $f(E^{n+1}) \subset X^*$, $f(\dot{E}^{n+1}) \subset X$, assuming that $f^{-1}(y_1 + \cdots + y_k) = f^{-1}(y_1)$, say, is a single circuit $\dot{E}^2$, where $E^2 \subset E^{n+1}$. The map f is to be simplicial in $f^{-1}(A_1^n)$, while E^2 is a sub-complex of some rectilinear subdivision of E^{n+1}. Therefore if the above simplex B_1^n is sufficiently small, it follows from a straightforward geometrical argument that, after starting again, $f^{-1}(\dot{A}_1^n)$ cuts E^2 in a simple circuit bounding a 2-element $E_0^2 \subset E^2$.

Let

$$K = E_0^2 + f^{-1}(\dot{A}_1^n).$$

I say that the identical map of $\dot{E}^{n+1}$ on itself is homotopic in $E^{n+1} - f^{-1}(a_1^n)$ to a map $u_2(\dot{E}^{n+1}) \subset K$. For the part of $f^{-1}(y_1)$ which lies in any $(n + 1)$-simplex $A^{n+1} \subset f^{-1}(A_1^n)$ is a linear segment, whose end points are internal to the two n-simplexes in $\dot{A}^{n+1}$ which cover A_1^n, whence it is clear that $f^{-1}(\dot{A}_1^n)$ is a retract by deformation of $f^{-1}(A_1^n) - f^{-1}(y_1)$. Let u be a regular neighborhood of K in $E^{n+1} - f^{-1}(a_1^n)$ and let $E_0^{n+1} \subset u + f^{-1}(a_1^n)$ be a regular neighborhood of E^2, which is internal to E^{n+1} (even if u meets $\dot{E}^{n+1}$), and which contains E^2 in its interior. Then E^{n+1} contracts into[23] E_0^{n+1}, and it follows that the identical map $u_0(\dot{E}^{n+1}) = \dot{E}^{n+1}$ is deformable in $E^{n+1} - f^{-1}(y_1)$ into a map $u_1(\dot{E}^{n+1}) = \dot{E}_0^{n+1} \subset u + \{f^{-1}(a_1^n) - f^{-1}(y_1)\}$. The complex K is a retract by deformation of u, and hence a retract by a deformation in which each point of K is held fixed,[24] and $f^{-1}(\dot{A}_1^n)$ is a retract by a similar deformation, relative to $f^{-1}(\dot{A}_1^n)$, of $f^{-1}(A_1^n) - f^{-1}(y_1)$. Therefore the map $u_1(\dot{E}^{n+1})$ can be deformed, in $E^{n+1} - f^{-1}(y_1)$, into a map $u_2(\dot{E}^{n+1}) \subset K$, by deformations of the two parts lying in $f^{-1}(A_1^n) - f^{-1}(y_1)$ and in u, with the common part, in $f^{-1}(\dot{A}_1^n)$, held fixed. Since $f^{-1}(\dot{A}_1^n)$ is a retract by deformation of $f^{-1}(A_1^n) - f^{-1}(y_1)$, it follows that $E^{n+1} - f^{-1}(a_1^n)$ is a retract by deformation of $E^{n+1} - f^{-1}(y_1)$. Therefore the deformation cylinder of the deformation $u_0 \to u_2$ may be deformed into $E^{n+1} - f^{-1}(a_1^n)$, and the assertion is justified. It follows that $f(\dot{E}^{n+1})$ is deformable, in X, into the map $g(\dot{E}^{n+1}) \subset X$, where $g = fu_2$.

[23] S.S., pp. 248, 258, 260 and Theorem 23, corollary 1 (p. 293).

[24] S.S., p. 273.

Since $n > 2$ the circuit $f(\dot{E}_0^2) \subset \dot{A}_1^n$ is deformable, in $\dot{A}_1^n$, into the point x_1. This deformation can be extended to a deformation, in $\dot{A}_1^n$, of $f\{f^{-1}(\dot{A}_1^n)\}$, and hence to a deformation $f_t(K) \subset X$ $(0 \leqq t \leqq 1; f_0 = f)$, such that $f_1(\dot{E}_0^2) = x_1$, $f_t\{f^{-1}(\dot{A}_1^n)\} = \dot{A}_1^n$. Therefore $f(\dot{E}^{n+1})$ is homotopic in X to $g_1(\dot{E}^{n+1})$, where $g_t = f_t u_2$. Let S^2 be a 2-sphere which meets $\dot{A}_1^n$ in the single point x_1, and let $h(\mathcal{E}_0^2) = S^2 - x_1$ be a homeomorphism of $\mathcal{E}_0^2 = E_0^2 - \dot{E}_0^2$ on $S^2 - x_1$. Let $h^*(K) = S^2 + \dot{A}_1^n$ be the map given by

$$h^* = h \quad \text{in} \quad \mathcal{E}_0^2$$
$$= f_1 \quad \text{in} \quad f^{-1}(\dot{A}_1^n),$$

and $f^*(S^2 + \dot{A}_1^n) \subset X$ the map given by

$$f^*(x) = f_1 h^{-1}(x) \quad \text{if} \quad x \epsilon S^2 - x_1$$
$$= x \quad \text{if} \quad x \epsilon \dot{A}_1^n.$$

Then $f_1(K) = f^*h^*(K)$, and

$$g_1(\dot{E}^{n+1}) = f^*g^*(\dot{E}^{n+1}) \subset X,$$

where $g^*(\dot{E}^{n+1}) = h^*u_2(\dot{E}^{n+1}) \subset S^2 + \dot{A}_1^n$. If we take x_1 as the base point for $\pi_n(S^2 + \dot{A}_1^n)$ it follows from theorem 2 that the element in $\pi_n(S^2 + \dot{A}_1^n)$, which is given by $g^*(\dot{E}^{n+1})$, is of the form $\beta_1 + \beta_2 + \alpha\cdot\beta$, where $\beta_1 \epsilon \pi_n(S^2)$, $\beta_2 \epsilon \pi_n(\dot{A}_1^n)$, $\alpha \epsilon \pi_{n-1}(\dot{A}_1^n)$ and $\beta \epsilon \pi_2(S^2)$. Clearly $g^*(\dot{E}^{n+1})$ is homotopic to a point in $S^2 + A_1^n$. Therefore $\beta_1 + \beta_2 + \alpha\cdot\beta$ reduces to zero if we fill in the simplex A_1^n. Since this has the same algebraic effect as mapping A_1^n on x_1 it follows that

$$\beta_1 = \phi_1(\beta_1 + \beta_2 + \alpha\cdot\beta) = 0,$$

where ϕ_1 means the same as in theorem 2, with S_1^m replaced by S^2 and S^n by $\dot{A}_1^n$. Let us transfer the base point of $\pi_n(X)$ to x_1, which is permissible since σ_n is an invariant sub-group, and let $\phi^*\{\pi_n(S^2 + \dot{A}_1^n)\} \subset \pi_n(X)$ be the homomorphism induced by the map $f^*(S^2 + \dot{A}_1^n) \subset X$. Then it follows from the definition of $\alpha\cdot\beta$ that $\phi^*(\alpha\cdot\beta) = \phi^*(\alpha)\cdot\phi^*(\beta)$ and hence that the element given by $f^*g^*(\dot{E}^{n+1})$ belongs to σ_n. This completes the proof.

Collecting these results we have the theorem:

THEOREM 3. *The residue group $\pi_n^* - \pi_n^0$ is isomorphic to $\mathfrak{M}_0$, and $\pi_n - \sigma_n$ is isomorphic to π_n^0. The homomorphism $\psi(\pi_n^*) = \mathfrak{M}_0$ determines an isomorphism $\psi(\pi_n^* - \pi_n^0) = \mathfrak{M}_0$ and $\psi(\pi_n) = \pi_n^0$ determines an isomorphism $\psi(\pi_n - \sigma_n) = \pi_n^0$.*

Notice that lemma 2 is valid, for obvious reasons, when $n \geqq 2$, provided we take $\pi_1 = \pi_1(X^*)$, which coincides with $\pi_1(X)$ if $n > 2$. Let $\tilde{X}^*$ be the universal covering space of X^* and let $\tilde{X}$ be the part of $\tilde{X}^*$ which covers X. By an adaption of Reidemeister's theory, referred to in §1 above, we may interpret m as the group of relative n-cycles in $\tilde{X}^*$ (mod $\tilde{X}$). Now let $n = 2$ and let m_0 be the sub-group of m, which consists of the absolute (singular) 2-cycles in $\tilde{X}^*$. Then $\psi(\alpha^*) \subset m_0$, and ψ is a homomorphism of π_2^* in m_0. If X and X^* are

simplicial complexes it follows from Hurewicz' theorem, referred to in §1, that $\psi(\pi_2^*) = m_0$.

6. Let Γ be a given group, whose elements we denote by small greek letters, let H be an aggregate of individuals, which we denote by $a, b, a_i, b_j, \cdots$, and let $h(H) \subset \Gamma$ be a single-valued, but not necessarily $(1-1)$, transformation of H in Γ. We shall use h_Γ to stand for the group generated by all the pairs (a, ξ), which we denote by a_ξ, subject to the relations

$$a_{\xi\alpha} = a_\xi, \qquad a_\xi b_\eta = b_\zeta a_\xi, \tag{6.1}$$

for each $a, b \subset H$ and $\xi, \eta \subset \Gamma$, where $\alpha = h(a)$ and $\zeta = \xi\alpha\xi^{-1}\eta$. From the first of these, and induction on $|n|$ $(n = 0, \pm1, \pm2, \cdots)$, we have

$$a_{\xi\alpha^n} = a_\xi, \tag{6.2}$$

and it may be verified that the second implies

$$a_\xi^\delta b_\eta^\epsilon = b_\zeta^\epsilon a_\xi^\delta \qquad (\delta, \epsilon = \pm 1), \tag{6.3}$$

where $\zeta = \xi\alpha^\delta\xi^{-1}\eta$, with $\alpha = h(a)$. If $x = a_\xi^\delta \cdots b_\eta^\epsilon$ is any element in h_Γ, and $\tau \in \Gamma$, let

$$\theta_\tau(x) = a_{\tau\xi}^\delta \cdots b_{\tau\eta}^\epsilon.$$

Since $\tau(\xi\alpha) = (\tau\xi)\alpha$ and $(\tau\xi\alpha\xi^{-1}\tau^{-1})(\tau\eta) = \tau(\xi\alpha\xi^{-1}\eta)$, the transformation given by $x \to \theta_\tau(x)$, for each product of generators, leaves the system of relations invariant. It therefore determines a homomorphism of h_Γ in itself. Clearly $\theta_{\tau^{-1}}\theta_\tau(x) = \theta_\tau\theta_{\tau^{-1}}(x) = x$, whence θ_τ is an automorphism. Also $\theta_\tau\theta_{\tau'} = \theta_{\tau\tau'}$, whence $\tau \to \theta_\tau$ is a homomorphism of Γ in the group of automorphisms of h_Γ.

Let $\phi(a_\xi^\epsilon) = \xi\alpha^\epsilon\xi^{-1}$ for any $a_\xi \in h_\Gamma$, where $\alpha = h(a)$. Since $\xi\alpha\alpha\alpha^{-1}\xi^{-1} = \xi\alpha\xi^{-1}$ and

$$\begin{aligned}\xi\alpha\xi^{-1}\eta\beta\eta^{-1} &= \xi\alpha\xi^{-1}\eta\beta\eta^{-1}\xi\alpha^{-1}\xi^{-1}\xi\alpha\xi^{-1} \\ &= \zeta\beta\zeta^{-1}\xi\alpha\xi^{-1},\end{aligned}$$

where $\beta = h(b)$ and $\zeta = \xi\alpha\xi^{-1}\eta$, the transformation given by

$$\phi(a_\xi^\delta \cdots b_\eta^\epsilon) = \phi(a_\xi^\delta) \cdots \phi(b_\eta^\epsilon)$$

is a homomorphism $\phi(h_\Gamma) \subset \Gamma$. Clearly $\phi(h_\Gamma)$ is the minimum invariant subgroup of Γ which contains $h(H)$. Notice also that $\phi\{\theta_\tau(x)\} = \tau\phi(x)\tau^{-1}$ for any $x \in h_\Gamma$, $\tau \in \Gamma$, whence $\theta_\tau\{\phi^{-1}(1)\} \subset \phi^{-1}(1)$. I say that $\phi^{-1}(1)$ is contained in the centrum of h_Γ. For let x and y be any elements in h_Γ, given by

$$x = a_{1\xi_1}^{\delta_1} \cdots a_{m\xi_m}^{\delta_m}, \qquad y = b_{1\eta_1}^{\epsilon_1} \cdots b_{n\eta_n}^{\epsilon_n}.$$

It follows from (6.3) and induction on $m + n$ that $xy = zx$, where

$$z = b_{1\zeta_1}^{\epsilon_1} \cdots b_{n\zeta_n}^{\epsilon_n}, \tag{6.4}$$

with[25] $\zeta_i = \phi(x)\eta_i$. Therefore $\phi(x) = 1$ implies $\zeta_i = \eta_i$, whence $z = y$. Let us now write the Abelian group $\phi^{-1}(1)$ with addition, and let us write $\theta_\tau(x) = \tau x$ for any $x \in \phi^{-1}(1)$. Then $\rho x \in \phi^{-1}(1)$ may be defined in the usual way, where ρ is any element in the integral group ring $\mathfrak{R}(\Gamma)$. Therefore $\phi^{-1}(1)$ is a group with operators in $\mathfrak{R}(\Gamma)$.

Returning to the point in §5 at which we required $n > 2$, we now take $n = 2$. Let $\alpha^* \in \pi_2^*$ be given by $f(S^2, p_0) \subset X^*$, which we assume to be simplicial in $f^{-1}(A_i^2)$. We shall simplify our notation by rewriting $s_{i\lambda}$, $p_{i\lambda}$, $\epsilon_{i\lambda}$ and $A_{i\lambda}^2 \subset S^2$ $(i = 1, \cdots, k;\ \lambda = 1, \cdots, q_i)$ as s_λ, p_λ, ϵ_λ and A_λ $(\lambda = 1, \cdots, q = q_1 + \cdots + q_k)$, where $f(A_\lambda) = \epsilon_\lambda A_{i_\lambda}^2$ and the segments $s_1, \cdots, s_q$ occur in this cyclic order round their common end point p_0. With a notation explained in S.S. p. 279, let $\Sigma = \Sigma_1 + \cdots + \Sigma_q$, where

$$\Sigma_\lambda = s_\lambda + \dot{A}_\lambda - s_\lambda,$$

and Σ_λ does not meet Σ_μ except at p_0, and s_λ is non-singular and does not meet A_λ except at p_λ. The singular circuit

$$f(\Sigma) = f(\Sigma_1) + \cdots + f(\Sigma_q) \subset X,$$

in which $f(s_1)$ is described first, represents the element

$$\xi = \xi_1 \alpha_{i_1}^{\epsilon_1} \xi_1^{-1} \cdots \xi_q \alpha_{i_q}^{\epsilon_q} \xi_q^{-1} \in \pi_1, \tag{6.5}$$

where α_i means the same as in §5 and ξ_λ is the element given by the circuit $f(s_\lambda) - l_{i_\lambda}$. If h_{π_1} and $\phi(h_{\pi_1}) \subset \pi_1$ mean the same as before, with $\Gamma = \pi_1$, $H = (a_1, \cdots, a_k)$ and $h(a_i) = \alpha_i$, we have

$$\xi = \phi\{\psi(f, \Sigma)\},$$

where ξ is given by (6.5), and

$$\psi(f, \Sigma) = a_{i_1 \xi_1}^{\epsilon_1} \cdots a_{i_q \xi_q}^{\epsilon_q}. \tag{6.6}$$

As in §5, the circuit $f(\Sigma)$ bounds the cell $f\left\{S^2 - \sum_{\lambda=1}^{q} (s_\lambda + A_\lambda)\right\}$. Therefore $\xi = 1$ and $\psi(f, \Sigma) \in \phi^{-1}(1)$. If $f(S^2) \subset X$ we set $\psi(f, \Sigma) = 0$.

Conversely, if x, given by the right hand side of (6.6), is an arbitrary element in h_{π_1}, we can construct $\Sigma \subset S^2$ and define a map $f(\Sigma) \subset X$, which represents the product $\xi = \phi(x)$, given by (6.5). If $\xi = 1$ the map $f(\Sigma)$ may be extended to a map $f(S^2) \subset X^*$ and we shall have $x = \psi(f, \Sigma)$. Therefore every element in $\phi^{-1}(1)$ is of the form $\psi(f, \Sigma)$ for a suitable choice of f and Σ.

Corresponding to lemmas 2 and 3 we have the theorem, in which π_n^0 $(n = 2)$ means the same as when $n > 2$:

Theorem 4. *The element $\psi(\alpha^*) \in \phi^{-1}(1)$, given by* (6.6), *depends only on the element $\alpha^* \in \pi_2^*$, given by $f(S^2, p_0)$. The transformation $\alpha^* \to \psi(\alpha^*)$ is an operator homomorphism of π_2^* on $\phi^{-1}(1)$, and $\psi^{-1}(0) = \pi_n^0$.*

---

[25] Notice the general form of the relations (6.3), namely $xyx^{-1} = \theta_{\phi(x)}(y)$.

In proving this we shall be concerned with $\phi^{-1}(1)$ as a sub-group of h_{π_1}, and shall therefore write it with multiplication. The first part of the theorem will follow from an argument used in lemma 2, when we have proved that $\psi(f, \Sigma) = 1$ if $\alpha^* = 0$. Therefore we assume that $S^2 = \dot{A}^3$, where A^3 is a rectilinear 3-simplex, and that there is a map $f(A^3) \subset X^*$. We also assume that f is simplicial in $f^{-1}(A_i^2)$. Let y_i be an inner point of A_i^2. Then, the trivial case $f(S^2) \subset X$ excepted, $f^{-1}(y_1 + \cdots + y_k) = L$, say, in a linkage which consists of non-singular, polygonal segments joining points on $\dot{A}^3$ and, possibly, circuits which are internal to A^3. We shall show that there is a homomorphism $F(G) \subset h_{\pi_1}$, where $G = \pi_1(A^3 - L)$, such that $\psi(f, \Sigma) = F(1) = 1$.

We take p_0 to be a vertex of A^3 and assume that the projection of L from p_0 on the opposite face is regular,[26] as the term is used in the theory of knots. Replacing $A_1^2, \cdots, A_k^2$ by smaller simplexes if necessary, we also assume that there are no accidental intersections in $f^{-1}(A_1^2 + \cdots + A_k^2)$; also that $T^2 = f^{-1}(\dot{A}_1^2 + \cdots + \dot{A}_k^2)$ cuts the cone swept out by the linear segment p_0p, as p varies over L, in a series of non-singular segments and circuits which are approximately parallel to the components of L. Let $L_1, \cdots, L_m$ be the segments of L which are 'completely visible' from p_0. That is to say the projection of L_ρ $(\rho = 1, \cdots, m)$ from p_0 does not pass under any segment of L, and each end point of L_ρ is either on $\dot{A}^3$, or is at a crossing of which L_ρ is a lower branch. Let $y_{i_\rho} = f(L_\rho)$, let $p \in L_\rho$, let p_ρ' be the point at which the rectilinear segment p_0p pierces T^2 and let p_ρ'' be a near-by point on T^2 such that $f(p_\rho'') = x_{i_\rho} \in \dot{A}_{i_\rho}^2$. Let c_ρ be a meridian circuit on T^2, beginning and ending at p_ρ'', which is oriented so that $f(c_\rho) = \dot{A}_{i_\rho}^2$, and let $l_\rho = p_0p_\rho' + p_\rho'p_\rho''$, where $p_0p_\rho' \subset p_0p$ and $p_\rho'p_\rho''$ is a segment on T^2 joining p_ρ' to p_ρ''. Then the group G is generated by $g_1, \cdots, g_m$, where g_ρ is given by the circuit $l_\rho + c_\rho + l_\rho$. We write

$$F(g_\rho) = a_{i_\rho\eta_\rho},$$

where $\eta_\rho \in \pi_1$ is given by the circuit $f(l_\rho) - t_{i_\rho}$. It follows from the same argument as when L is an ordinary knot or linkage that the relations determined by the crossings in the projection constitute a complete set.[27] Let

$$g_\rho^\epsilon g_\lambda = g_\mu g_\rho^\epsilon \tag{6.7}$$

be such a relation (see the diagram, in which $\epsilon = 1$).

The circuit $l_\rho + \epsilon c_\rho - l_\rho$ is obviously homotopic, rel. p_0, to a circuit of the form $l_\mu + p_\mu p_\lambda - l_\lambda$, where $p_\mu p_\lambda$ is a segment on T^2 which joins p_μ to p_λ. Since $f(p_\mu p_\lambda) \subset \dot{A}_{i_\mu}^2$ $(i_\mu = i_\lambda)$, the circuit $t_{i_\mu} + f(p_\mu p_\lambda) - t_{i_\lambda}$ represents an element of the form $\alpha_{i_\mu}^n$. Therefore $f(l_\mu + p_\mu p_\lambda - l_\lambda)$, which is homotopic, rel. p_0, to

$$\{f(l_\mu) - t_{i_\mu}\} + \{t_{i_\mu} + f(p_\mu p_\lambda) - t_{i_\lambda}\} - \{f(l_\lambda) - t_{i_\lambda}\},$$

---

[26] See, for example, K. Reidemeister, Knotentheorie, Berlin (1932), 5.

[27] See Knotentheorie, p. 54. This fact is, so to speak, half the content of the proof. For it follows from an argument given below that $\psi(f, \Sigma) = 1$ is equivalent to the relation corresponding to $\dot{A}^3$, when the latter is treated a single multiple point of the graph L.

represents the element $\eta_\mu\alpha^{n}_{i_\mu}\eta_\lambda^{-1}$. Similarly $f(l_\rho + \epsilon c_\rho - l_\rho)$ represents the element $\eta_\rho\alpha^{\epsilon}_{i_\rho}\eta_\rho^{-1}$. Therefore

$$(6.8) \qquad \eta_\mu\alpha^{n}_{i_\mu} = \eta_\rho\alpha^{\epsilon}_{i_\rho}\eta_\rho^{-1}\eta_\lambda .$$

From (6.3) we have

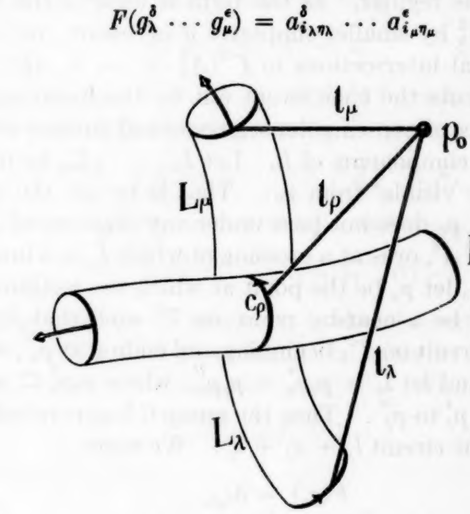
$$a^{\epsilon}_{i_\rho\eta_\rho}a_{i_\lambda\eta_\lambda} = a_{i_\lambda\zeta}a^{\epsilon}_{i_\rho\eta_\rho} ,$$

where $\zeta = \eta_\rho\alpha^{\epsilon}_{i_\rho}\eta_\rho^{-1}\eta_\lambda$, and it follows from (6.8) and (6.2) that

$$(6.9) \qquad a^{\epsilon}_{i_\rho\eta_\rho}a_{i_\lambda\eta_\lambda} = a_{i_\mu\eta_\mu}a^{\epsilon}_{i_\rho\eta_\rho} .$$

On comparing (6.7) with (6.9), we see that the transformation given by

$$F(g^{\delta}_{\lambda} \cdots g^{\epsilon}_{\mu}) = a^{\delta}_{i_\lambda\eta_\lambda} \cdots a^{\epsilon}_{i_\mu\eta_\mu}$$



is a homomorphism $F(G) \subset h_{\pi_1}$. Let $\phi_f(G) \subset \pi_1$ be the homomorphism of G determined by the map f and the base point $f(p_0) = x_0$. Then $\phi_f(g_\rho) = \eta_\rho\alpha_{i_\rho}\eta_\rho^{-1}$, whence $\phi_f = \phi F$.

Let c be a meridian circuit on T^2, oriented so that $f(c) = \dot{A}^2_i$ for some value of i, and let $l \subset A^3 - L$ be any segment which joins p_0 to a point $p \epsilon f^{-1}(x_i)\cdot c$. Let $g \epsilon G$ be the element given by $l + c - l$, and let $\xi \epsilon \pi_1$ be the element given by $f(l) - t_i$. I say that

$$(6.10) \qquad F(g) = a_{i\xi} .$$

For let pp''_ρ be a segment on T^2 which joins p to some p''_ρ. Then g is also given by $(l + pp''_\rho) + c_\rho - (l + pp''_\rho)$. On the other hand

$$f(l + pp''_\rho) - t_i = f(l) + f(pp''_\rho) - t_i ,$$

which is homotopic, rel p_0 , to

$$\{f(l) - t_i\} + \{t_i + f(pp''_\rho) - t_i\},$$

and since $f(pp_\rho'') \subset \mathring{A}_i^2$ it follows that the element of π_1 which is given by $f(l + pp_\rho'') - t_i$ is of the form $\xi\alpha_i^n = \eta$, say, whence $a_{i\xi} = a_{i\eta}$, in consequence of (6.2). Therefore we may replace c by c and l by $l + pp_\rho''$ without altering either g of $a_{i\xi}$. To simplify the notation we shall start again, assuming that $c = c_p$, $p = p_\rho''$. Then $g = \bar{g}g_\rho\bar{g}^{-1}$, where $\bar{g}$ is the element given by $l - l_\rho$, and

$$F(g) = \bar{x}a_{i\eta_\rho}\bar{x}^{-1} \qquad (i = i_\rho),$$

where $\bar{x} = F(\bar{g})$. It follows from (6.4) that

$$F(g) = a_{i\xi},$$

where $\xi = \phi(\bar{g})\eta_\rho$. But $\phi(\bar{x}) = \phi F(\bar{g}) = \phi_f(\bar{g})$, and since $\bar{g}$ is given by $l - l_\rho$ the element $\phi(\bar{x})$ is given by $f(l) - f(l_\rho)$. Therefore $\xi = \phi(\bar{x})\eta_\rho$ is given by $f(l) - f(l_\rho) + f(l_\rho) - t_i$, or by $f(l) - t_i$, as stated.

It follows from the preceding paragraph that $F(g^*) = \psi(f, \Sigma)$, where $g^* \in G$ is given by the circuit Σ. But $g^* = 1$, since Σ bounds a cell in $\mathring{A}^3$. Therefore $\psi(f, \Sigma) = 1$, and the first part of the proof is complete.

The fact that ψ is a homomorphism follows from the argument used to reduce the first part to the case where $\alpha^* = 0$. The argument used when $n > 2$ shows that ψ is an operator homomorphism.

It follows from the definition that $\psi(\pi_2^0) = 1$, the group $\phi^{-1}(1)$ still being multiplicative. Conversely let $\psi(\alpha^*) = 1$ and let α^* be given by a map $f(S^2) \subset X^*$ which is simplicial in $f^{-1}(A_1^2 + \cdots + A_k^2)$. Then the product (6.6), determined by the map f and some $\Sigma \subset S^2$, represents the identity in h_{π_1}. Therefore it can be transformed into a product of the form

$$x_1R_1^{\epsilon_1}x_1^{-1} \cdots x_nR_n^{\epsilon_n}x_n^{-1}, \tag{6.11}$$

where each R_λ is of the form

$$a_{i\xi\alpha_i}a_{i\xi}^{-1} \quad \text{or} \quad a_{i\xi}a_{j\eta}a_{i\zeta}^{-1}a_{j\zeta}^{-1} \qquad (\zeta = \xi\alpha_i\xi^{-1}\eta),$$

by a sequence of operations which consist of cancelling, or of inserting, consecutive terms of the form $a_{i\xi}^{\epsilon}a_{i\xi}^{-\epsilon}$. Each cancelling operation can be copied by the geometrical process described in lemma 3, which can obviously be reversed so as to copy an insertion. Therefore we may assume in the first place that the product (6.6) is of the form (6.11). Moreover we may subject the factors $a_{i\xi}^{\epsilon}$ of (6.6) and, at the same time, the subscripts $1, \cdots, q$ in $\Sigma = \Sigma_1 + \cdots + \Sigma_q$, to any cyclic permutation. Therefore, after transferring x_1, from the beginning to the end of (6.11), and inserting $x_1x_1^{-1}$ between x_λ^{-1} and $x_{\lambda+1}$, for each $\lambda = 2, \cdots, n - 1$, we may further assume that $x_1 = 1$.

The geometrical process by which terms of the form xx^{-1} are removed involves an alteration of the map f. We now leave f as it is, but replace $s_1, \cdots, s_q$ by segments $s_1', \cdots, s_q'$, in such a way as to transform $\psi(f, \Sigma)$, given by (6.1), into

$$\psi(f, \Sigma') = Wx_2R_2^{\epsilon_2}x_2^{-1} \cdots x_nR_n^{\epsilon_n}x_n^{-1},$$

where W is of the form aa^{-1} or $aa^{-1}bb^{-1}$.

First let $R_1 = a_{i\xi\alpha_i}a_{i\xi}^{-1}$. Then $R_1^{-1} = a_{i\eta\alpha_i^{-1}}a_{i\eta}^{-1}$, where $\eta = \xi\alpha_i$, and replacing ξ by $\xi\alpha_i^{-1}$ if $\epsilon = -1$, we have $R_1^\epsilon = a_{i\xi\alpha_i^\epsilon}a_{i\xi}^{-1}$. Let

$$\Sigma' = s_1' + \dot{A}_1' - s_1' + \cdots + s_q' + \dot{A}_q' - s_q',$$

where $A_\lambda' = A_\lambda$ ($\lambda = 1, \cdots, q$) and

$$s_1' = s_1 - \epsilon\dot{A}_1 \quad \text{(S.D.)}, \qquad s_\lambda' = s_\lambda \qquad (\lambda = 2, \cdots, q),$$

in which (S.D.) means that the segment indicated is to be slightly deformed so as to eliminate accidental intersections. Then, remembering that $f(A_1) = A_i^2$, we see that $f(s_1') - t_i$ is homotopic, rel. p_0, to

$$f(s_1) - t_i + (t_i - \epsilon\dot{A}_i^2 - t_i).$$

Since $f(s_1) - t_i$ represents the element $\xi\alpha_i^\epsilon$, it follows that the circuit $f(s_1') - t_i$ represents the element $\xi\alpha_i^\epsilon\alpha_i^{-\epsilon} = \xi$. Therefore

$$\psi(f, \Sigma') = a_{i\xi}a_{i\xi}^{-1}x_2R_2^{\epsilon_2}x_2^{-1} \cdots x_nR_n^{\epsilon_n}x_n^{-1}. \tag{6.12}$$

Secondly let $R_1 = a_{i\xi}a_{j\eta}a_{i\xi}^{-1}a_{j\zeta}^{-1}$, where $\zeta = \xi\alpha_i\xi^{-1}\eta$. Then $R_1^{-1} = a_{j\zeta}a_{i\xi}a_{j\eta}^{-1}a_{i\xi}^{-1}$, and if $\epsilon_1 = -1$ we replace every factor $x_\lambda R_\lambda^{\epsilon_\lambda}x_\lambda^{-1}$ in (6.11) by $a_{j\zeta}^{-1}x_\lambda R_\lambda^{\epsilon_\lambda}x_\lambda^{-1}a_{j\zeta}$ and cancel the first two terms $a_{j\zeta}^{-1}a_{j\zeta}$. This operation can be copied geometrically, and the result is to replace R_1^{-1} by $a_{i\xi}a_{j\eta}^{-1}a_{i\xi}^{-1}a_{j\zeta}$. Therefore we may take $R_1^\epsilon = a_{i\xi}a_{j\eta}^\epsilon a_{i\xi}^{-1}a_{j\zeta}^{-\epsilon}$, Let

$$\Sigma' = s_1' + \dot{A}_1' - s_1' + \cdots + s_q' + \dot{A}_q' - s_q',$$

where

$$s_1' = s_1 + \dot{A}_1 - s_1 + s_2 \quad \text{(S.D.)}, \qquad s_2' = s_1, \qquad s_\lambda' = s_\lambda,$$

$$A_1' = A_2, \qquad A_1' = A_1, \qquad A_\lambda' = A_\lambda \qquad (\lambda = 3, \cdots, q).$$

Then $f(A_1') = A_j^2$, $f(A_2') = A_i^2$ and $f(s_1') - t_j$ is homotopic to

$$\{f(s_1) - t_i\} + (t_i + \dot{A}_i^2 - t_i) - \{f(s_1) - t_i\} + \{f(s_2) - t_j\},$$

and therefore represents the element $\xi\alpha_i\xi^{-1}\eta = \zeta$. The circuit $f(s_2') - t_i = f(s_1) - t_i$ represents the element ξ, and we have

$$\psi(f, \Sigma') = a_{j\zeta}^\epsilon a_{i\xi}a_{i\xi}^{-1}a_{j\zeta}^{-\epsilon}x_2R_2^{\epsilon_2}x_2^{-1} \cdots x_nR^{\epsilon_n}x_n^{-1}. \tag{6.13}$$

If $n = 1$ it follows from (6.12) or (6.13) that $f(S^2)$ is homotopic to a map in X. In general, the terms preceding x_2, in (6.12) or in (6.13), can be removed by cancelling, and it follows from induction on n that $\alpha^* \epsilon \pi_2^0$. Therefore π_2^0 is the kernel of ψ and the theorem is established.

COROLLARY. *If $\pi_2(X) = 0$, then $\pi_2(X^*)$ is isomorphic to $\phi^{-1}(1)$ under the transformation ψ.*

As an application of this corollary, let X^* be a 2-dimensional complex and let $\mathfrak{E}_1^n, \cdots, \mathfrak{E}_k^n$ be all the 2-cells in X^*, with the notation used at the beginning of §5. Then X_0 is the linear graph which consists of all the 1-cells in X^*. There-

fore $\pi_2(X) = 0$ and $\pi_1(X)$ is a free group. Therefore, given a product $a_\xi^\delta \cdots b_\eta^\epsilon$, of the generators of h_{π_1}, we can decide whether or no $\phi(a_\xi^\delta \cdots b_\eta^\epsilon) = 1$. Let F be the free group, which is freely generated by the generators of h_{π_1}, and let $\phi(x) \epsilon \pi_1(X)$ have the obvious meaning if $x \epsilon F$. Then the (multiplicative) group π_2^* can be calculated in the form:

Generators: *all the elements of* $\phi^{-1}(1) \subset F$,
Relations: *all relations of the form* $x = y$, *where* $x, y \subset \phi^{-1}(1)$ *are of the form* $x = x_1x_2$, $y = x_1Rx_2$, *in which* R *is of the form* $a_{\xi\alpha}a_\xi^{-1}$ or $a_\xi b_\eta a_\xi^{-1} b_\zeta^{-1}$ $(\alpha = \phi(a), \zeta = \xi\alpha\xi^{-1}\eta)$.

By this method we not only calculate π_2^* as an abstract group but, given a map $f(S^2, p_0) \subset X^*$, with $f(p_0) = x_0$, we can calculate $\psi(\alpha^*) \epsilon \phi^{-1}(1)$, where α^* is the element represented by $f(S^2, p_0)$. Conversely, given $x \epsilon \phi^{-1}(1)$ we can construct a map which represents $\psi^{-1}(x) \epsilon \pi_2^*$, as in the preamble to theorem 4. Therefore, given any simplicial complex K, we can first calculate $\pi_2(K^2)$ and then $\pi_2(K^3) = \pi_2(K)$, by means of S. S., theorem 18, where K^n is the n-dimensional skeleton of K.

If X^* is a polyhedron the kernel of the homomorphism $\chi(\pi_2) = \pi_2^0$ can be expressed in terms of Reidemeister's[28] theory of homology with coefficients in $\mathfrak{R}$, or in the residue ring $\mathfrak{R} - \mathfrak{J}$, where $\mathfrak{J}$ is any two-sided ideal in $\mathfrak{R}$. For let $\tilde{X}^*$ be the universal covering space of X^*, and let $\bar{X} = u^{-1}(X)$, where $u(\tilde{X}^*) = X^*$ is a regular covering of X^* by $\tilde{X}^*$. Let $\tilde{X}$ be the universal covering space of $\bar{X}$, and hence of X, and let $\bar{u}(\tilde{X}) = \bar{X}$ be a regular covering of $\bar{X}$ by $\tilde{X}$. Then[29]

$$\pi_2(X) \simeq \pi_2(\tilde{X}) \simeq \beta_2(\tilde{X}), \qquad \pi_2(X^*) \simeq \pi_2(\tilde{X}^*) \simeq \beta_2(\tilde{X}^*),$$

where $\simeq$ denotes isomorphism and $\beta_2(P)$ is the second homology group of P, with integral coefficients. More precisely, $\pi_2(\tilde{X})$ is isomorphic to $\beta_2(\tilde{X})$ in the transformation under which $\alpha \epsilon \pi_2$, given by $f(S^2) \subset \tilde{X}$, corresponds to the homology class containing the cycle $f(S^2)$. The homomorphism $\chi(\pi_2) = \pi_2^0$ is the one determined by the map $\bar{u}(\tilde{X}) \subset \tilde{X}^*$. It follows that the homology classes corresponding to the elements in $\chi^{-1}(0)$ are those which reduce to zero in consequence of the relations $\alpha_1 = \cdots = \alpha_k = 0$.

In case X is a finite, 2-dimensional complex one can also express $\chi^{-1}(0)$ as follows. Let $X = X_1 = K^1 + B_1^2 + \cdots + B_m^2$, $X^* = X + A_1^2 + \cdots + A_k^2$, and let $X_2 = K^1 + A_1^2 + \cdots + A_k^2$, where K^1 is a linear graph. Let us rewrite the group h_{π_1}, which is determined by X and X^*, as $h(X, X^*)$. Then $h(K^1, X^*)$ is generated by the generators $a_{i\xi}$ $(i = 1, \cdots, k)$ of $h(K^1, X_1)$, together with the generators $b_{j\eta}$ $(j = 1, \cdots, m)$ of $h(K^1, X_2)$, where $\xi, \eta \subset F = \pi_1(K^1)$. The relations for $h(K^1, X^*)$ consist of the relations $R_{\rho\lambda}$, for $h(K^1, X_\rho)$ $(\rho = 1, 2;$ $\lambda = 1, 2, \cdots)$, together with a system of relations $R_{12\lambda}$, which are of the form

$$a_{i\xi}b_{j\eta} = b_{j\zeta}a_{i\xi}.$$

[28] Loc. cit.

[29] Hurewicz (loc. cit.), paper II, p. 522.

Then $\chi^{-1}(0)$ is isomorphic to the sub-group of $\phi^{-1}(1) \subset h(K^1, X_1)$, whose elements reduce to 1 on the introduction of the new generators $b_{j\eta}$ and the additional relations $R_{2\lambda}$ and $R_{12\lambda}$. In particular $\pi_2 = \pi_2^0$ if every relation between the generators $a_{i\xi}$ is a consequence of the relations $R_{1\lambda}$.

By way of application consider the question: *Is any sub-complex of an aspherical,*[30] *2-dimensional complex itself aspherical?* Let $K^2 = K^1 + A_1^2 + \cdots + A_k^2 + B_1^2 + \cdots + B_m^2$ be any 2-dimensional complex, let $K_1^2 = K^1 + B_1^2 + \cdots + B_m^2$, $K_2^2 = K^1 + A_1^2 + \cdots + A_k^2$ and let us denote arbitrary elements of $h(K^1, K_1^2)$, of $h(K^1, K_2^2)$ and of $h(K^1, K^2)$ by x, by y and by z or z' respectively. We recall the general form of the relations for $h(K^1, K^2)$ which follows from (6.4), namely $zz'z^{-1} = g_z(z')$, where $g_z = \theta_{\phi(z)}$. Since $z \to \phi(z)$ and $\zeta \to \theta_\zeta$ $(\zeta \in \pi_1)$ are homomorphisms, it follows that $z \to g_z$ is a homomorphism of $h(K^1, K^2)$ in its group of automorphisms. The group $h(K^1, K^2)$ is obtained from the free product $h(K^1, K_1^2) \circ h(K^1, K_2^2)$ by adjoining the relations $xyx^{-1} = g_x(y)$, $yxy^{-1} = g_y(x)$, for every $x \in h(K^1, K_1^2)$, $y \in h(K^1, K_2^2)$. Thus the above question is part of a wider question, which can be stated as follows. Let G_1 and G_2 be given groups, and to each $x \in G_1$, $y \in G_2$, let there correspond automorphisms $y' = f_x(y)$, $x' = g_y(x)$, of G_2 and G_1, such that the transformations $x \to f_x$ and $y \to g_y$ are homomorphisms of G_1 and G_2 in the automorphic groups of G_2 and G_1. Then the question is: *Under what conditions are G_1 and G_2, regarded as sub-groups of $G_1 \circ G_2$, unaffected by the additional relations $xyx^{-1} = f_x(y)$, $yxy^{-1} = g_y(x)$, in which x and y range over all the elements in G_1 and G_2?*

BALLIOL COLLEGE
OXFORD, ENGLAND.

[30] Hurewicz (loc. cit.), paper IV. See also J. H. C. Whitehead, Fund. Math., 32 (1939), 149–66.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

ON HUMBERT FUNCTIONS

BY R. S. VARMA

(Received December 12, 1939)

I. INTRODUCTION

The function $J_{m,n}(x)$ defined by the relation

$$(1) \qquad J_{m,n}(x) = \frac{(x/3)^{m+n}}{\Gamma(m+1)\Gamma(n+1)} {}_0F_2\left(m+1, n+1; -\frac{x^3}{27}\right)$$

was first studied by P. Humbert[1] in the year 1930 and later on, among other results, he gave the following operational relations:[2]

$$(2) \qquad x^{\frac{2m-n}{3}} J_{m,n}(3\sqrt[3]{x}) \doteq \frac{1}{p^{\frac{2m-n}{2}}} J_n\left(-2\sqrt{\frac{1}{p}}\right)$$

and

$$(3) \qquad x^{\frac{2n-m}{3}} J_{m,n}(3\sqrt[3]{x}) \doteq \frac{1}{p^{\frac{2n-m}{2}}} J_m\left(-2\sqrt{\frac{1}{p}}\right).$$

The function $J_{m,n}(x)$ has been called by Humbert a Bessel function of the third kind, but in order to avoid confusion with the ordinary Bessel function of the third kind, I shall call it a Humbert function.

The object of this paper is to investigate[3] some properties of Humbert functions so far as the convergence of infinite series involving the functions are concerned. In general we can have the following two types of infinite series involving the functions:

$$(A) \qquad \sum_{n=1}^{\infty} A_{m,n} J_{m,n}(x)$$

and

$$(B) \qquad \sum_{n=1}^{\infty} A_{ln+k,n} J_{ln+k,n}(x).$$

[1] P. Humbert: Les fonctions de Bessel du troisieme ordre, Atti. Pont. Acad. della Scienza, Anno. LXXXIII (Sess. III del 16 Febbraio, 1930), 128–146.

[2] P. Humbert: Nouvelles remarques sur les fonctions de Bessel du troisieme ordre, ibid, Anno. LXXXVII (Sess. IV del 18 Marzo, 1934), 323–331.

[3] The discussion regarding the asymptotic behavior of $J_{m,n}(x)$ for large x will follow in a separate communication to this Journal.

The convergence of these two types of infinite series are discussed in §§2–3. An operational relation between a Humbert function and a Kummer function ${}_1F_1$ is deduced in §4. In the subsequent article an infinite series involving the product of a Humbert function and a Weber's parabolic cylinder function $D_n(x)$ is summed up in terms of Kelvin's function bei(x). Finally in §6 the summation of an infinite series involving a Humbert function and a Neumann's polynomial $0_n(t)$ are effected by means of parabolic cylinder functions.

II. Series of the Type (A)

From the relation (1), it is evident that for large n,

$$J_{m,n}(x) = \frac{(x/3)^n}{\Gamma(n+1)}[1 + O(n^{-1})]. \tag{4}$$

This by virtue of Stirling's formula

$$\Gamma(n) = n^{n-\frac{1}{2}}e^{-n}(2\pi)^{\frac{1}{2}}e^{\theta/12n} \tag{5}$$

gives that

$$J_{m,n}(x) = O\left(\frac{(x/3)^n}{n^{n+\frac{1}{2}}e^{-n}}\right) \tag{6}$$

for large n.

The series (A) is therefore convergent throughout the x-domain in which

$$\overline{\lim_{n\to\infty}} \sqrt[n]{\left|\left\{A_{m,n}\frac{(x/3)^n}{e^{-n}n^{n+\frac{1}{2}}}\right\}\right|} < 1.$$

In case the series (A) occurs in the form

$$\sum_{n=1}^{\infty} A_{m,n}\left(\frac{x}{3}\right)^{-n} J_{m,n}(x) \tag{A$_1$}$$

the necessary and sufficient condition for the convergence of (A_1) is that

$$\left|\frac{A_{m,n}}{e^{-n}n^{n-\frac{1}{2}}}\right| \tag{A$_1'$}$$

should tend to zero as n tends to infinity.

Example 1. The relation (1) can be written as

$$J_{m,n}(x) = \frac{(x/3)^{m+n}}{\Gamma(m+1)}\sum_{r=0}^{\infty}\frac{(-x^3/27)^r}{r!(m+1,r)\Gamma(n+r+1)}$$

where

$$(l, r) = l(l+1)(l+2)\cdots(l+r-1) = \frac{\Gamma(l+r)}{\Gamma(l)}.$$

This gives that

$$\sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!} \left(\frac{x}{3}\right)^{-n} J_{m,n}(x)$$

$$= \frac{(x/3)^m}{\Gamma(m+1)} \sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!} \sum_{r=0}^{\infty} \frac{(+x^3/27)^r}{(r!)^2(m+1, r)(r+1, n)}$$

$$= \frac{(x/3)^m}{\Gamma(m+1)} \sum_{r=0}^{\infty} \frac{(-x^3/27)^r}{(r!)^2(m+1, r)} \sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!(r+1, n)}$$

the inversion of the order of summation being justified on account of the absolute convergence of the two series involved.

But we know that

$$\sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!(r+1, n)} = {}_2F_1(\alpha, \beta; r+1; 1)$$

$$= \frac{\Gamma(r+1)\Gamma(r+1-\alpha-\beta)}{\Gamma(r+1-\alpha)\Gamma(r+1-\beta)}$$

$$R(r+1-\alpha-\beta) > 0.$$

Hence

$$\sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!} \left(\frac{x}{3}\right)^{-n} J_{m,n}(x)$$

$$= \frac{(x/3)^m \Gamma(1-\alpha-\beta)}{\Gamma(1+m)\Gamma(1-\alpha)\Gamma(1-\beta)} \sum_{r=0}^{\infty} \frac{(1-\alpha-\beta, r)(-x^3/27)^r}{r!(m+1, r)(1-\alpha, r)(1-\beta, r)} \tag{7}$$

$$= \frac{(x/3)^m \Gamma(1-\alpha-\beta)}{\Gamma(1+m)\Gamma(1-\alpha)\Gamma(1-\beta)} {}_1F_3\left[1-\alpha-\beta; 1+m, 1-\alpha, 1-\beta; \frac{x^3}{27}\right]$$

provided that $R(\alpha+\beta) < 1$.

The series (7) is convergent, since the condition (A_1') here reduces after a little algebra, to $n^{\alpha+\beta-1}$, which on account of the condition $R(\alpha+\beta) < 1$, tends to zero as n tends to infinity.

The generalized hypergeometric series on the right of (7) reduces to Humbert functions when $m = -\alpha - \beta$ and when $\beta = 0$. Hence we obtain as special cases of (7), the following relations:

$$\sum_{n=0}^{\infty} \frac{(\alpha, n)(\beta, n)}{n!} \left(\frac{x}{3}\right)^{-n} J_{-\alpha-\beta,n}(x) = J_{-\alpha,-\beta}(x) \qquad R(\alpha+\beta) < 1$$

and

$$\sum_{n=0}^{\infty} \frac{(\alpha, n)}{n!} \left(\frac{x}{3}\right)^{-n} J_{m,n}(x) = \left(\frac{x}{3}\right)^{\alpha} \Gamma(1-\alpha) J_{m,-\alpha}(x) \qquad R(\alpha) < 1.$$

EXAMPLE 2. Another example of the series of the type (A) is the infinite series

$$\sum_{n=0}^{\infty} \frac{J_{m,n}(-3x^2/4)}{n!} = (-x^2/4)^m \sum_{r=0}^{\infty} \frac{(x/2)^{5r}}{r!\Gamma(m+r+1)} J_r(x), \tag{8}$$

which can be easily established by the help of the definition (1), remembering also that

$$J_r(x) = \frac{(x/2)^r}{r!} {}_0F_1(r+1; -\tfrac{1}{4}x^2).$$

III. SERIES OF THE TYPE (B)

When $m = ln + k$, the series (1) reduces to

$$J_{ln+k,n}(x) = \frac{(x/3)^{ln+n+k}}{\Gamma(ln+n+k)\Gamma(k+1)} {}_0F_2\left(ln+k+1, n+1; -\frac{x^3}{27}\right).$$

This with the help of the estimate (5) gives that, for large n,

$$J_{ln+k,n}(x) = O\left(\frac{(x/3)^{ln+n}}{e^{-ln-n}\, l^{ln}\, n^{n+ln+k+1}}\right).$$

The series of the type (B) therefore converges throughout the x-domain in which

$$\overline{\lim_{n\to\infty}} \sqrt[n]{\left|\left\{A_{ln+k,n} \frac{(x/3)^{ln+n}}{e^{-ln-n}\, l^{ln}\, n^{n+ln+k+1}}\right\}\right|} < 1.$$

In case the series (B) is of the form

$$\sum_{n=1}^{\infty} A_{ln+k,n} \left(\frac{x}{3}\right)^{-ln-n} J_{ln+k,n}(x) \tag{B$_1$}$$

the necessary and sufficient condition for the convergence will be that

$$\left|\frac{A_{ln+k,n}}{e^{-ln-n}\, l^{ln}\, n^{n+ln+k}}\right| \tag{B$_1'$}$$

should tend to zero as n tends to infinity.

EXAMPLE 1. Humbert[4] has investigated four series of the type. Of these we mention here

$$\sum_{n=0}^{\infty} \frac{(kx)^n}{3^n n!} J_{n,n}(x) = J_{o,o}(x\sqrt[3]{1+k})$$

from which, by putting $k = -1$, he has obtained the series

$$\sum_{n=0}^{\infty} \frac{(-x/3)^n}{n!} J_{n,n}(x) = 1.$$

---

[4] P. Humbert: Second paper quoted above, pp. 326–327. We can easily establish the convergence of the series investigated by Humbert by the help of the estimate of $J_{m,n}(x)$ given in this paper. There seems to be an omission of sign by Humbert in the series quoted here.

I propose to give now another generalization of this particular series by proving that

$$\sum_{n=0}^{\infty} \frac{(-x/3)^n}{n!} J_{n+\frac{1}{2}\nu, n+\frac{1}{2}\nu}(x) = \frac{(-)^{\frac{1}{2}\nu}(x/3)^\nu}{[\Gamma(\frac{1}{2}\nu + 1)]^2} \tag{9}$$

which for $\nu = 0$ reduces to Humbert's series just given above.

To establish (9), consider the known series[5]

$$\sum_{n=0}^{\infty} \frac{(\frac{1}{2}z)^{n+\frac{1}{2}\nu}}{n!} J_{n+\frac{1}{2}\nu}(z) = \frac{(z/2)^\nu}{\Gamma(\frac{1}{2}\nu + 1)}.$$

In this put $z = -2/\sqrt{p}$ where

$$\frac{1}{p\alpha} \doteqdot \frac{x^\alpha}{\Gamma(\alpha + 1)} \qquad R(\alpha + 1) > 0 \tag{10}$$

we then get

$$\sum_{n=0}^{\infty} \frac{(-)^n}{n!\, p^{\frac{1}{2}n+\frac{1}{2}\nu}} J_{n+\frac{1}{2}\nu}\left(-\frac{2}{\sqrt{p}}\right) = \frac{(-)^{\frac{1}{2}\nu}}{p^{\frac{1}{2}\nu}\Gamma(\frac{1}{2}\nu + 1)}.$$

Interpreting the left hand side by (2) and the right hand side by (10), then Lerch's theorem gives that

$$\sum_{n=0}^{\infty} \frac{(-)^n}{n!} x^{\frac{1}{2}n+\frac{1}{2}\nu} J_{n+\frac{1}{2}\nu, n+\frac{1}{2}\nu}(3\sqrt[3]{x}) = \frac{(-x)^{\frac{1}{2}\nu}}{[\Gamma(\frac{1}{2}\nu + 1)]^2}$$

which can be thrown in the form (9).

EXAMPLE 2. An interesting series of the type (*B*) is given by

$$\sum_{r=0}^{\infty} \frac{(-\frac{1}{2}\sin 2\theta)^{\nu+2r}}{r!\,\Gamma(\nu + r + 1)} x^{\frac{1}{2}(\nu+2r)} J_{\nu+2r,\nu+2r}(3\sqrt[3]{x}) = J_{\frac{1}{2}\nu,\nu}(3\sqrt[3]{x\cos^2\theta}) J_{\frac{1}{2}\nu,\nu}(3\sqrt[3]{x\sin^2\theta}).$$

This series is important in as much as it gives the expansion of the product of two Humbert functions of different arguments in an infinite series involving Humbert functions.

Now if we put $m = \frac{1}{2}\nu$ and $n = \nu$, the image (2) reduces to

$$J_{\frac{1}{2}\nu,\nu}(3\sqrt[3]{x}) \doteqdot J_\nu(-2/\sqrt{p}).$$

If we use this and follow the method of example 1, this series can be easily established by interpreting both sides of the known series[6]

$$\sum_{r=0}^{\infty} \frac{(\frac{1}{4}z\sin 2\theta)^{\nu+2r}}{r!\,\Gamma(\nu + r + 1)} J_{\nu+2r}(z) = J_\nu(z\cos\theta) J_\nu(z\sin\theta).$$

A special case of the series investigated above is given by putting $\theta = \pi/4$, when we get that

$$\sum_{r=0}^{\infty} \frac{(-\frac{1}{2})^{\nu+2r}}{r!\,\Gamma(\nu + r + 1)} x^{\frac{1}{2}(\nu+2r)} J_{\nu+2r,\nu+2r}(3\sqrt[3]{x}) = [J_{\frac{1}{2}\nu,\nu}(3\sqrt[3]{x/2})]^2.$$

[5] Watson: Bessel Function, p. 525.

[6] W. N. Bailey: On the product of two Legendre Polynomials with different arguments, Proc. Lond. Math. Soc., (2), 41 (1936), 215–220.

IV. An Operational Image for Humbert Functions

A theorem[7] of operational calculus is that if

$$f(x) \doteqdot \phi(p)$$

then

$$f(x^2) \doteqdot \frac{p}{\sqrt{\pi}} \int_0^\infty e^{-\frac{1}{4}p^2s^2} \phi\left(\frac{1}{s^2}\right) ds. \tag{11}$$

Taking

$$f(x) = x^{\frac{2m-n}{3}} J_{m,n}(3\sqrt[3]{x})$$

and

$$\phi(p) = \frac{1}{p^{\frac{2m-n}{2}}} J_n\left(-2\sqrt{\frac{1}{p}}\right)$$

as given by (2), we get that

$$x^{\frac{4m-2n}{3}} J_{m,n}(3x^{2/3}) \doteqdot \frac{p}{\sqrt{\pi}} \int_0^\infty s^{2m-n} e^{-\frac{1}{4}p^2s^2} J_n(-2s)\, ds.$$

Since[8]

$$\int_0^\infty t^\lambda e^{-at^2} J_n(bt)\, dt = \frac{b^n \Gamma(\frac{1}{2}n + \frac{1}{2}\lambda + \frac{1}{2})}{2^{n+1} a^{\frac{1}{2}n+\frac{1}{2}\lambda+\frac{1}{2}} \Gamma(n+1)}$$

$$\times {}_1F_1\left(\tfrac{1}{2}n + \tfrac{1}{2}\lambda + \tfrac{1}{2}; n+1; -\frac{b^2}{4a^2}\right), \quad R(n + \lambda + 1) > 0.$$

It follows that

$$x^{\frac{4m-2n}{3}} J_{m,m}(3x^{2/3}) \doteqdot \frac{(-)^n (4/p^2)^m \Gamma(m + \frac{1}{2})}{\sqrt{\pi}\Gamma(n+1)} {}_1F_1\left(m + \tfrac{1}{2}; n+1; -\frac{4}{p^2}\right) \tag{12}$$

$$R(m + \tfrac{1}{2}) > 0.$$

Similarly if we use the relation (3) in (11), we obtain that

$$x^{\frac{4n-2m}{3}} J_{m,n}(3x^{2/3}) \doteqdot \frac{(-)^m (4/p^2)^n \Gamma(n + \frac{1}{2})}{\sqrt{\pi}\Gamma(m+1)} \tag{13}$$

$$\times {}_1F_1\left(n + \tfrac{1}{2}; m+1; -\frac{4}{p^2}\right).$$

The result (13) is also a consequence of (12), since $J_{m,n}(x)$ is symmetrical in m and n.

---

[7] P. Humbert: Le Calcul Symbolique (Paris, 1934), p. 28.

[8] Watson: Bessel Functions, p. 393.

3.2. We shall now show that there exist operational relations between Humbert functions and the various types of confluent hypergeometric functions.

Thus if we use Kummer's first transformation formula, viz.,

$$ {}_1F_1(\alpha; \rho; z) = e^z\, {}_1F_1(\rho - \alpha; \rho; -z) $$

(12) becomes

$$ x^{\frac{4m-2n}{3}} J_{m,n}(3x^{2/3}) \doteq \frac{(-)^n (4/p^2)^m e^{-4/p^2} \Gamma(m + \frac{1}{2})}{\sqrt{\pi}\Gamma(n+1)} \times {}_1F_1\left(n - m + \tfrac{1}{2}; n + 1; \frac{4}{p^2}\right). \tag{14} $$

Since Whittaker's function $M_{k,m}(x)$ is defined by the relation

$$ M_{k,m}(x) = x^{m+\frac{1}{2}} e^{-\frac{1}{2}x}\, {}_1F_1(\tfrac{1}{2} + m - k; 2m + 1; x) $$

(14) gives as a particular case, that

$$ x^{4/3r} J_{s+r,2s}(3x^{2/3}) \doteq \frac{(-)^s (4/p^2)^{r-\frac{1}{2}} \Gamma(s + r + \frac{1}{2})}{\sqrt{\pi}\Gamma(2s+1)} M_{r,s}(4/p^2) \quad R(s + r + \tfrac{1}{2}) > 0. $$

Taking appropriate values of m and n, we can similarly show that Humbert functions are operationally related to Laguerre polynomials $L_n^m(x)$, to Weber's parabolic cylinder functions $D_n(x)$, Bateman's functions $k_{2n}(x)$, and Bessel functions of the second kind $I_n(x)$.

V. An Infinite Series Involving the Product of a Humbert Function and a Parabolic Cylinder Function

If we put $b = \dfrac{2\sqrt{2}}{p}$ and $t = y\sqrt{2}$ in the known series[9]

$$ e^{-\frac{1}{4}t^2} \sin bt = e^{-\frac{1}{2}b^2} \sum_{n=0}^{\infty} \frac{(-)^n b^{2n+1}}{(2n+1)!} D_{2n+1}(t), $$

we get

$$ e^{-\frac{1}{2}y^2} \sin \frac{4y}{p} = \sum_{n=0}^{\infty} \frac{(-)^n 2^{3n+\frac{3}{2}} e^{-4/p^2}}{p^{2n+1}(2n+1)!} D_{2n+1}(y\sqrt{2}). \tag{15} $$

Consider p as a symbolic operator given by the relation (10). Putting $m = n + \frac{1}{2}$ in (12), we get that

$$ x^{\frac{2n+2}{3}} J_{n+\frac{1}{2},n}(3x^{2/3}) \doteq \frac{(-)^n}{\sqrt{\pi}} e^{-4/p^2} (4/p^2)^{n+\frac{1}{2}}. \tag{16} $$

If we now use (16) and the known operational image[10]

$$ \operatorname{bei}(2\sqrt{x}) \doteq \sin \frac{1}{p} $$

[9] R. S. Varma: On functions associated with the parabolic cylinder in harmonic analysis, Proc. Benares Math. Soc., 10 (1928), 15.

[10] Balth van der Pol: On the operational solution of linear differential equation and an investigation of the properties of these solutions, Phil. Mag. 8 (1929), 861–898.

and obtain the originals of both sides of (15), then Lerch's theorem gives that

$$\sum_{n=0}^{\infty} \frac{2^n}{(2n+1)!} x^{2n/3} J_{n+\frac{1}{3},n}(3x^{2/3}) D_{2n+1}(y\sqrt{2}) \tag{17}$$
$$= \frac{1}{\sqrt{(2\pi)}} x^{-2/3} e^{-\frac{1}{2}y^2} \operatorname{bei}(4\sqrt{xy}).$$

Since for large integral values[11] of n,

$$D_n(x) = \sqrt{2}(\sqrt{n})^n e^{-\frac{1}{2}n} \left[\cos(xn^{\frac{1}{2}} - \tfrac{1}{2}n\pi) + \frac{w_n(x)}{\sqrt{n}} \right]$$

where $w^n(x)$ satisfies both the inequalities

$$|w_n(x)| < \frac{3\cdot 35\ldots}{|x|\sqrt{\pi}} e^{\frac{1}{4}x^2}, \qquad |w_n(o)| < \tfrac{1}{6}n^{-\frac{1}{3}}$$

it is easy to see that the series (17) certainly converges, for all finite values of y within and on the circle $|x| = 1$.

VI. An Infinite Series Involving the Product of a Humbert Function and a Neumann's Polynomial

We consider the expansion

$$O_0(t)J_0(z) + 2\sum_{n=1}^{\infty} O_n(t)J_n(z) = \frac{1}{t-z} \qquad |z| < |t|.$$

Putting $z = -1/\sqrt{2p}$ and finding, in the manner of §5, the original of either side by the help of (10) and the known result[12]

$$(2x)^{\frac{1}{2}(m-1)} e^{\frac{1}{2}x} D_{-m}\{\sqrt{(2x)}\} \doteqdot \left(\frac{\pi}{2}\right)^{\frac{1}{2}} \frac{\sqrt{p}}{(\sqrt{p}+1)^m} \qquad R(m) > -1$$

we get that

$$O_0(t)J_{0,0}(\tfrac{3}{2}\sqrt[3]{x}) + 2\sum_{n=1}^{\infty} O_n(t) J_{\frac{1}{3}n,n}(\tfrac{3}{2}\sqrt[3]{x}) \tag{18}$$
$$= \left(\frac{\pi}{2}\right)^{-\frac{1}{2}} t^{-1} e^{x/4t^2} D_{-1}(\sqrt{x}/t).$$

If we use the relation[13]

$$O_n(t) = \frac{2^{n-1}\, n!}{t^{n+1}} \{1 + \phi_n\}$$

where $\phi_n \to 0$ as $n \to \infty$, we find that the region of convergence is confined to the domain given by

$$\overline{\lim_{n\to\infty}} \sqrt[n]{\left| \left\{ \left(\frac{\sqrt{x}}{t}\right)^n \frac{e^{\frac{1}{2}n}}{n^{\frac{1}{2}n+\frac{1}{2}}} \right\} \right|} < 1.$$

Lucknow, India.

[11] Whittaker and Watson: Modern Analysis (Fourth Edition), p. 354.
[12] R. S. Varma: Summation of some infinite series of Weber's parabolic cylinder functions, Jour. Lond. Math. Soc., 12 (1937), 25–27.
[13] Whittaker and Watson: Modern Analysis (Fourth Edition), p. 375.

Annals of Mathematics
Vol. 42, No. 2, April, 1941

ON TWO PROBLEMS OF SAMPLING

By Brockway McMillan

(Received January 29, 1940)

I

We consider two problems which can be stated in the following terms: we have an arbitrarily large number of urns, each containing a mixture of white and black balls. If p_k is the density of white balls in the k^{th} urn (probability of drawing a white ball from the k^{th} urn), we suppose that the *a priori* probability that $p_k < x$ is given by $f_k(x)$, where

$$\begin{aligned} &f_k(x) = 0, \qquad x \leqq 0; \\ &f_k(x) = 1, \qquad x > 1; \\ &f_k(x) \leqq f_k(x+h), \qquad h \geqq 0; \\ &\lim_{h\to 0+} f_k(x-h) = f_k(x). \end{aligned} \qquad k = 1, 2, \cdots \tag{1}$$

We suppose that these probabilities for separate urns are independent. We seek by taking a sample drawing of one ball from each of the first N urns to investigate the average density

$$\frac{p_1 + p_2 + \cdots p_N}{N}$$

of white balls in these urns. Our two different problems arise from two different methods of treating the sample.

For the asymptotic results which we wish to state, it will be necessary to think of the number of urns as being infinite, and to think of the sample as being taken from the whole array. The results of a drawing will be specified by an infinite sequence of indices, $(z_k, k = 1, 2, \cdots)$, defined by

$$\begin{aligned} z_k &= 1 \text{ if white is drawn from the } k^{\text{th}} \text{ urn,} \\ z_k &= 0 \text{ otherwise, } k = 1, 2, \cdots \end{aligned}$$

First problem: We suppose that the sample $(z_k, k = 1, 2, \cdots, N)$ is known exactly, and ask the *a posteriori* probability $Q_N((z_k), x)$ that

$$\frac{p_1 + p_2 + \cdots p_N}{N} < x. \tag{2}$$

Second problem: We suppose only that the density m/N of white balls in the sample is known, $m = z_1 + z_2 + \cdots z_N$ and ask the *a posteriori* probability $P_N(m/N, x)$ of (2).

We investigate the asymptotic behavior of $Q_N((z_k), x)$ and $P_N(m/N, x)$ for large N. When the functions $f_k(x)$ are all the same function, the analytical formulations of these two problems become identical. This simpler case has been treated by Bochner[1] and v. Mises[2]. Their results carry over to the more general problems when the moments of the $f_k(x)$ are sufficiently restricted. We have, in fact, under conditions to be stated below, that both $Q_N((z_k), x)$ and $P_N(m/N, x)$ tend with increasing N to a function $P(x)$ with the properties

$$(3)\qquad \begin{aligned} P(x) &= 0, \qquad x < p_0, \\ P(x) &= 1, \qquad x > p_0, \end{aligned}$$

where the "probable density" p_0 is defined in terms of the sample and the moments of the $f_k(x)$.

We suppose first that the moments

$$(4)\qquad a_k = \int_0^1 x\, df_k(x), \qquad k = 1, 2, \cdots,$$

are neither zero nor unity:

$$(5)\qquad a_k(1 - a_k) \neq 0, \qquad k = 1, 2, \cdots .$$

We next state a series of definitions, each for $k = 1, 2, \cdots$.

$$(6)\qquad \begin{aligned} g_k(x) &= \frac{1}{a_k} \int_{-\infty}^{x} p\, df_k(p) \\ h_k(x) &= \frac{1}{1 - a_k} \int_{-\infty}^{x} (1 - p)\, df_k(p) \end{aligned}$$

$$(7)\qquad \begin{aligned} b_{n,k} &= \int_0^1 x^n\, dg_k(x) \\ c_{n,k} &= \int_0^1 x^n\, dh_k(x) \end{aligned} \qquad n = 0, 1, 2, \cdots$$

$$(8)\qquad b_k = b_{1,k}, \qquad c_k = c_{1,k}$$

$$(9)\qquad d_k = b_{2,k} - (b_k)^2, \qquad e_k = c_{2,k} - (c_k)^2.$$

We note that the functions $g_k(x)$, $h_k(x)$ are distribution functions satisfying conditions of the form of (1). From this it follows that their dispersions (9) satisfy

$$(10)\qquad 0 \leqq d_k \leqq 1, \qquad 0 \leqq e_k \leqq 1, \qquad k = 1, 2, \cdots .$$

We can now make more complete statements of the results indicated above. We recall that (5) is a restriction already imposed upon the $f_k(x)$.

[1] S. Bochner, "A Converse of Poisson's Theorem." Ann. Math. 37, 1936, pp. 816–822.
[2] R. v. Mises, "A Modification of Bayes' Problem." Ann. Math. Stat. 9, 1938, pp. 256–259.

(A) The sample (z_k) defines a sequence of distribution functions

$$F_k(x) = z_k g_k(x) + (1 - z_k)h_k(x), \qquad k = 1, 2, \cdots, \tag{11}$$

whose means r_k and dispersions s_k are given by

$$\begin{aligned} r_k &= z_k b_k + (1 - z_k)c_k \\ s_k &= z_k d_k + (1 - z_k)e_k . \end{aligned} \qquad k = 1, 2, \cdots \tag{12}$$

If the series $s_1 + s_2 + \cdots$ diverges, and if the sequence (r_k) has an average,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} r_k = r,$$

then $Q_N((z_k), x)$ tends as $N \to \infty$ to $P(x)$ of (3), with $p_0 = r$.

(B) Here we separate the restrictions on the sample (z_k) from those on the functions $f_k(x)$. We assume that the sample has a density:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} z_k = z. \tag{13}$$

Defining

$$l_k = \min. (d_k , e_k), \qquad k = 1, 2, \cdots, \tag{14}$$

we assume that the series $l_1 + l_2 + \cdots$ diverges. We further suppose that the sequences (b_k) and (c_k) possess averages in a certain strong sense: i.e., that there exist constants b and c such that[3]

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} (b_k - b)^2 = 0, \\ &\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} (c_k - c)^2 = 0. \end{aligned} \tag{15}$$

Under these hypotheses, both $Q_N((z_k), x)$ and $P_N(m/N, x)$ tend to $P(x)$ of (3) with

$$p_0 = zb + (1 - z)c.$$

The strict asymptotic form of these results, as above stated, may be proved by an application of the Laplace-Liapounoff limit theorem, in the way that v. Mises[2] treats the simpler problem. We shall use the more powerful methods of Bochner[1], since they enable us to give estimates of the errors involved in replacing $Q_N((z_k), x)$ and $P_N(m/N, x)$ by $P(x)$. We shall, in fact, state our results without reference to limits as $N \to \infty$, in forms which will be valid for finite samples.

---

[3] Because the sequences in question are bounded, these are equivalent to conditions of the form $\lim N^{-1} \sum | b_k - b | = 0$.

II

This section reproduces essentially the argument of Bochner.[1] We assume (5). We consider a fixed sequence of zeros and ones, the sample (z_k, $k = 1, 2, \cdots, N$). Nearly all quantities with which we deal will depend upon the sample considered, but for simplicity of notation we shall not in this section attempt to indicate explicitly the fact of this dependence.

Let us assume for the moment that the sample is

$$(16)\qquad \begin{aligned} z_k &= 1, \qquad k = 1, 2, \cdots, m; \\ z_k &= 0, \qquad k = m+1, \cdots, N. \end{aligned}$$

By Bayes' theorem, $Q_N((z_k), x) = Q_N(x)$ is given by

$$Q_N(x) = \frac{J_N(x)}{J_N(1)}$$

where $J_N(x)$ is the compound probability that simultaneously (2) shall hold and the drawing yield (16):

$$J_N(x) = \int_{p_1+p_2+\cdots+p_N<Nx} \cdots \int p_1 p_2 \cdots p_m(1 - p_{m+1}) \cdots (1 - p_N)\, df_1(p_1) \cdots df_N(p_N).$$

$J_N(1)$ can be calculated immediately. Recalling (4), we have

$$(17)\qquad J_N(1) = a_1 a_2 \cdots a_m(1 - a_{m+1}) \cdots (1 - a_N).$$

In terms of the $F_k(x)$ defined by (11), $J_N(x)$ may be written

$$(18)\qquad J_N(x) = a_1 \cdots a_m(1 - a_{m+1}) \cdots (1 - a_N) \int_{p_1+\cdots+p_N<Nx} \cdots \int dF_1(p_1) \cdots dF_N(p_N),$$

from which, by (17), we have

$$(19)\qquad Q_N(x) = \int_{p_1+\cdots+p_N<Nx} \cdots \int dF_1(p_1) \cdots dF_N(p_N).$$

This last clearly holds independently of the temporary assumption (16).

It is known (Bochner[1]) that (19) may be expressed in the form

$$(20)\qquad Q_N(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} E_N(w) \frac{\sin Nxw}{w}\, dw$$

where

$$E_N(w) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\,(iw(p_1 + p_2 + \cdots p_N))\, dF_1(p_1) \cdots dF_N(p_N).$$

This becomes

$$E_N(w) = \prod_{k=1}^{N} G_k(w)$$

where

$$G_k(w) = \int_0^1 e^{iwp}\, dF_k(p).$$

Consider the functions $\exp(-iwr_k)G_k(w)$, where r_k is defined by (12). Because

$$e^{-iwr_k}G_k(w) = \int_0^1 e^{iw(p-r_k)}\, dF_k(p),$$

we have the expansion

$$e^{-iwr_k}G_k(w) = \sum_{n=0}^{\infty} \frac{(iw)^n}{n!} G_{n,k}$$

valid for all w, with

$$G_{n,k} = \int_0^1 (x - r_k)^n\, dF_k(x), \qquad n = 0, 1, 2, \cdots, k = 1, 2, \cdots.$$

In particular, referring to (12), $G_{0,k} = 1$, $G_{1,k} = 0$, $G_{2,k} = s_k$. Furthermore, since $0 \leqq r_k \leqq 1$, for $n = 1, 2, \cdots$ we have

$$0 \leqq |G_{2+n,k}| \leqq \int_0^1 |x - r_k|^{2+n}\, dF_k(x) \leqq \int_0^1 |x - r_k|^2 dF_k(x)$$
$$= G_{2,k} = s_k \leqq 1.$$

Therefore

$$e^{-iwr_k}G_k(w) = 1 - s_k \frac{w^2}{2} + R_k(w) \tag{21}$$

where for $|w| \leqq 1$

$$|R_k(w)| \leqq s_k |w|^3 e. \tag{22}$$

By (10) we have then, uniformly in k

$$|R_k(w)| \leqq 3 |w|^3. \tag{23}$$

Defining

$$D(N) = s_1 + s_2 + \cdots + s_N,$$

we have from (21) and (22), for some w_0, $0 < w_0 \leqq 1$, and all w with $|w| \leqq w_0$, that

$$\log [e^{-iw(r_1+r_2+\cdots r_N)} E_N(w)] = -D(N) \frac{w^2}{2} + 0(w^3 D(N)).$$

From this, (23), and the identity

$$|U - V| = |U^{1/N} - V^{1/N}| \cdot |U^{(N-1)/N} + U^{(N-2)/N}V^{1/N} + \cdots V^{(N-1)/N}|$$

follows the existence of constants $A > 0$, $B > 0$, such that

$$| E_N(w)e^{-iw(r_1+r_2+\cdots r_N)} - e^{-\frac{1}{2}D(N)w^2} | \leqq AN \, | w |^3 e^{-BD(N)w^2} \tag{24}$$

holds for all w, $| w | \leqq w_0$, and all N.

Since we know only that $| E_N(w) | \leqq 1$, in place of (20) we consider the absolutely convergent integral

$$I_N(x) = \int_0^x Q_N(p)\,dp = \frac{2}{\pi}\int_{-\infty}^{\infty} E_N(w)\,\frac{(\sin \frac{1}{2}Nxw)^2}{Nw^2}\,dw.$$

The estimate (24) enables us to conclude that $I_N(x)$ differs from

$$\frac{2}{\pi}\int_{-\infty}^{\infty} [e^{iw(r_1+r_2+\cdots r_N)-\frac{1}{2}D(N)w^2}]\,\frac{(\sin \frac{1}{2}Nxw)^2}{Nw^2}\,dw \tag{25}$$

by

$$0\left[\int_{-w_0}^{w_0} | w | \, e^{-BD(N)w^2}\,dw\right] + 0\left[\int_{w_0}^{\infty} \frac{dw}{Nw^2}\right] = 0\left(\frac{1}{D(N)}\right) + 0\left(\frac{1}{N}\right).$$

Defining

$$r(N) = \frac{1}{N}(r_1 + r_2 + \cdots + r_N), \qquad H(N) = \frac{D(N)}{2N^2},$$

we rewrite (25) with $-w/N$ in place of w:

$$\frac{2}{\pi}\int_{-\infty}^{\infty} [e^{-iwr(N)-w^2H(N)}]\,\frac{(\sin \frac{1}{2}xw)^2}{w^2}\,dw. \tag{26}$$

This is equal to

$$\frac{1}{2\pi}\int_0^x dt \int_{-t}^{t} du \int_{-\infty}^{\infty} e^{iw(u-r(N))-w^2H(N)}\,dw$$

or

$$\frac{1}{2\sqrt{\pi H(N)}}\int_0^x dt \int_{-t}^{t} \exp\left(-\frac{(u - r(N))^2}{4H(N)}\right) du,$$

which becomes by integration by parts

$$\frac{x - r(N)}{2\sqrt{\pi H(N)}}\int_{-x-r(N)}^{x-r(N)} \exp\left(-\frac{u^2}{4H(N)}\right) du$$
$$+ \sqrt{\frac{H(N)}{\pi}}\left[\exp\left(-\frac{(x - r(N))^2}{4H(N)}\right) - \exp\left(-\frac{(x - r(N))^2}{4H(N)}\right)\right].$$

This differs from

$$\frac{x - r(N)}{2\sqrt{\pi H(N)}}\int_{-\infty}^{x-r(N)} \exp\left(-\frac{u^2}{4H(N)}\right) du$$

by a term which is majorized for $x \geqq 0$ by

$$0(\sqrt{H(N)}) + \frac{x - r(N)}{\sqrt{\pi}} \int_{\frac{x+r(N)}{2\sqrt{H(N)}}}^{\infty} e^{-u^2}\,du = 0(\sqrt{H(N)}) = 0\left(\frac{\sqrt{D(N)}}{N}\right).$$

In terms of the error function

$$\psi(x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{x} e^{-\frac{1}{4}u^2}\,du$$

we have altogether that for $x \geqq 0$

$$\left| I_N(x) - [x - r(N)]\cdot\psi\left[\frac{\sqrt{2N}(x - r(N))}{\sqrt{D(N)}}\right]\right| \leqq 0\left(\frac{\sqrt{D(N)}}{N}\right) + 0\left(\frac{1}{D(N)}\right). \tag{27}$$

We may use this to deduce result (A) stated above. By (10) we have that $D(N) \leqq N$. Defining

$$U(N) = | r - r(N) |,$$

we have from (27) that for $0 \leqq x \leqq \min(0, r - U(N))$

$$\int_0^x Q_N(p)\,dp \leqq 0\left(\frac{\sqrt{D(N)}}{N}\right) + 0\left(\frac{1}{D(N)}\right), \tag{29a}$$

while for $\min(1, r + U(N)) \leqq x \leqq 1$,

$$\left| \int_0^x Q_N(p)\,dp - (x - r)\right| \leqq U(N) + 0\left(\frac{\sqrt{D(N)}}{N}\right) + 0\left(\frac{1}{D(N)}\right). \tag{29b}$$

These and (10) imply that for $h > U(N)$

$$\begin{aligned} Q_N(r - h) &\leqq \frac{1}{h}\left[0\left(\frac{1}{\sqrt{N}}\right) + 0\left(\frac{1}{D(N)}\right)\right] \\ | Q_N(r + h) - 1 | &\leqq \frac{1}{h}\left[U(N) + 0\left(\frac{1}{\sqrt{N}}\right) + 0\left(\frac{1}{D(N)}\right)\right], \end{aligned} \tag{30}$$

since $Q_N(x)$ is non-decreasing in x. This establishes result (A) when $r = \lim r(N)$.

Using (24), with further restrictions on the $f_k(x)$, we may state an estimate analogous to (27) for $Q_N(x)$ itself. Essentially, all we require is that

$$\left| \int_{w_0}^{\infty} + \int_{-\infty}^{-w_0} E_N(w)\,\frac{\sin Nxw}{w}\,dw \right| \tag{31}$$

vanish as $N \to \infty$, for any x, $0 \leqq x \leqq 1$, and any w_0, $0 < w_0 \leqq 1$. If, for example, the distribution functions $f_k(x)$ have densities $f_k'(x)$ of variations bounded uniformly in k, (31) is dominated by $0(N^{-1})$. We suppose that for

any $w_0 > 0$ there is a function $V(N)$ which dominates (31) uniformly in $0 \leqq x \leqq 1$. We have by (24) then that $Q_N(x)$ differs from

$$\frac{1}{\pi}\int_{-\infty}^{\infty}[e^{iwNr(N)-\frac{1}{2}D(N)w^2}]\frac{\sin Nxw}{w}\,dw \tag{32}$$

by a term which is dominated by

$$0\left[N\int_{-w_0}^{w_0} w^2 e^{-BD(N)w^2}\,dw\right] + V(N) + 0\left[\int_{w_0}^{\infty} e^{-\frac{1}{2}D(N)w^2}\frac{dw}{w}\right]$$
$$= 0(ND(N)^{-\frac{3}{2}}) + V(N).$$

The integral (32) can be written

$$\frac{1}{2\pi}\int_{-x}^{x} du \int_{-\infty}^{\infty} e^{iw(u-r(N))-H(N)w^2}\,dw$$

or

$$\frac{1}{2\sqrt{\pi H(N)}}\int_{-x-r(N)}^{x-r(N)} \exp\left(-\frac{u^2}{4H(N)}\right)du.$$

If $r(N) \geqq \epsilon > 0$ for all N, this last differs from

$$\psi\left[\frac{\sqrt{2}N(x - r(N))}{\sqrt{D(N)}}\right]$$

by

$$0\left[\int_{-\infty}^{-\frac{\epsilon}{\sqrt{H(N)}}} e^{-\frac{1}{4}u^2}\,du\right] \leqq 0[\sqrt{H(N)}e^{-\frac{1}{4}\epsilon^2/H(N)}]$$

uniformly in $0 \leqq x \leqq 1$. Under the assumption, then, that $r(N) \geqq \epsilon > 0$ for large N, we have

$$\left|Q_N(x) - \psi\left[\frac{\sqrt{2}N(x - r(N))}{\sqrt{D(N)}}\right]\right| \leqq V(N) + 0(ND(N)^{-\frac{3}{2}}). \tag{33}$$

III

We now turn to our second problem. Again we assume (5). For convenience, we shall speak of a sequence $(z_k, k = 1, 2, \cdots, N)$ such that $z_1 + z_2 + \cdots + z_N = m$ as an m-sequence. We seek the probability $P_N(m/N, x)$ of (2) when it is known that an m-sequence has been drawn. By Bayes' theorem

$$P_N(m/N, x) = \frac{K_{m,N}(x)}{K_{m,N}(1)}$$

where $K_{m,N}(x)$ is the probability of the drawing of an m-sequence simultaneously with the occurrence of (2). That is, using the notation of the last section,

$$K_{m,N}(x) = \sum J_N((z_k), x) = \sum J_N((z_k), 1)Q_N((z_k), x)$$

where the sum is taken over all the C_m^N possible m-sequences. We have, therefore,

$$P_N(m/N, x) = \frac{\sum J_N((z_k), 1)Q_N((z_k), x)}{\sum J_N((z_k), 1)},$$

which exhibits $P_N(m/N, x)$ as a weighted average over all distinct m-sequences of the functions $Q_N((z_k), x)$ studied in the preceeding section. The results of that section will carry over, then to $P_N(m/N, x)$ once certain uniformity requirements have been met.[4]

Recalling (14), we define

$$L(N) = \sum_{k=1}^{N} l_k \leqq \min_{z_1+z_2+\cdots+z_N=m} [D((z_k), N)]$$

and

$$T(N) = \max_{z_1+z_2+\cdots z_N=m} | p_0 - r((z_k), N) |.$$

We have from (30) that for $h \geqq T(N)$

$$P_N(m/N, p_0 - h) \leqq \frac{1}{h} [0(N^{-\frac{1}{2}}) + 0(L(N)^{-1})]$$

(34)

$$| P_N(m/N, p_0 + h) - 1 | \leqq \frac{1}{h} [T(N) + 0(N^{-\frac{1}{2}}) + 0(L(N)^{-1})].$$

This, together with (30), will establish result (B). For any m-sequence and $p_0 = zb + (1 - z)c$ (see (13), (15))

$$\begin{aligned} |r((z_k), N) - p_0| &= \left| \frac{1}{N} \sum_{k=1}^{N} [z_k(b_k - b) + (1 - z_k)(c_k - c)] \right. \\ &\qquad\qquad \left. + \frac{m}{N} b + \left(1 - \frac{m}{N}\right) c - p_0 \right| \\ &\leqq \frac{1}{N} \sum_{k=1}^{N} [| b_k - b | + | c_k - c |] + \left| \frac{m}{N} - z \right| (b + c) \\ &\leqq \left[\frac{1}{N} \sum_{k=1}^{N} (b_k - b)^2 \right]^{\frac{1}{2}} + \left[\frac{1}{N} \sum_{k=1}^{N} (c_k - c)^2 \right]^{\frac{1}{2}} + 2 \left| \frac{m}{N} - z \right|, \end{aligned}$$

the last step by virtue of the Schwarz inequality. From this, we see that the hypotheses of (B) insure that $T(N) \to 0$; since they also insure that $L(N) \to \infty$, the result follows from (30) and (34).

An estimate anlogous to (33) for $P_N(m/N, x)$ is not possible without a restriction on the rate at which the limit in (13) is approached, and consequently has no meaning for finite samples. In this connection, it might be pointed out that the use of the error function $\psi(x)$ in stating, for example, inequalities (27) and (33) is largely a concession to convention. The right members of these inequalities would remain unchanged if $\psi(x)$ were replaced by a step function analogous to $P(x)$ of (3).

PRINCETON UNIVERSITY, PRINCETON, N. J.

[4] It is, in fact, for this reason that our first problem was introduced. All our attempts to apply directly to $P_N(m/N, x)$ analysis analogous to that in section II above bogged down in a morass of combinatory coefficients.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

TRANSFORMATIONS OF FINITE PERIOD. III

Newman's Theorem

BY P. A. SMITH

(Received February 22, 1940)

We shall give a new proof of a theorem of M. H. A. Newman [3] which asserts that periodic transformations of given period p operating in a locally euclidean space M can not be arbitrarily "small" (where by a small transformation is meant one which displaces points by a uniformly small amount). Our proof is quite different from Newman's and in a sense more direct; for only the space M which is being transformed comes into consideration whereas Newman's proof has for its setting the topological product of p copies of M. The methods which we shall use enable us to dispense with the locally euclidean restriction and to obtain a proof for far more general spaces. Moreover we establish a theorem which is slightly stronger than Newman's since it asserts, for periodic transformations operating in M, that there exists an impossible degree of smallness which is independent of the period. In this result, however, smallness must be understood in terms of orbits rather than displacements. To these considerations which in themselves perhaps justify a new proof of a known theorem, we add the following. The question whether or not there can exist *any* group—not merely finite, cyclic—operating effectively in a reasonably regular space and defining uniformly small orbits, is an outstanding problem in the theory of topological groups[1] and new methods of treating the special case under consideration will perhaps not be without interest.

A SPECIAL CASE

1. In order to make clear the nature of the theorem to be proved and to reveal the underlying ideas of the proof, which might otherwise remain hidden by the details, we shall first consider a simple example.

Let T be a transformation operating in a space M,—that is a one-one bicontinuous transformation of M into itself. The totality of images of a given point x of M under positive and negative powers of T constitute the *orbit* of x defined by T.

Let M_2 be an ordinary 2-sphere in euclidean 3-space. We shall call a *cap* of M_2 any closed circular region on M_2 which is smaller than a hemisphere. The special theorem to be proved is the following:

Every periodic transformation operating in M_2 defines at least one orbit which is not contained in any cap.

[1] For example, a negative answer to this question would establish the existence of one-parameter subgroups in locally euclidean topological groups.

Proof. Suppose on the contrary that there exists a transformation T of period q operating in M_2 such that every orbit defined by T is contained in a cap. It is easy to see that for each point x of M_2 there will be a uniquely determined smallest cap $u(x)$ containing the orbit of x defined by T and that the center $\omega(x)$ of $u(x)$ will be a continuous function of x.

We shall make use of the well known fact that every periodic transformation operating in a sphere is topologically equivalent to a periodic orthogonal transformation. From this it follows that there can be introduced into M_2 a triangulation which is preserved by T and whose 2-cells can be represented without repetition by

$$E^i, TE^i, \cdots, T^{q-1}E^i \qquad (i = 1, \cdots, \alpha).$$

We may assume that each E^i is positively oriented relative to a definitely chosen orientation of M. Let us assume for definiteness that T preserves orientation. Then the chain

$$\Delta = \sum_i (E^i + TE^i + \cdots + T^{q-1}E^i)$$

is, relative to every coefficient group, a fundamental 2-cycle on M_2 and as such, can not be ~ 0 on M_2. In particular Δ can not be $\sim 0 \bmod q$.

Let μ denote the single-valued continuous mapping $x \to \omega(x)$ defined over M_2; it carries Δ into a singular 2-cycle Δ^* on M. Moreover μ can be obtained by performing a deformation: we have merely to slide x to $\omega(x)$ along that geodesic arc which is shorter than a great semicircle, and at a suitable rate of speed depending on x. It follows that $\Delta \sim \Delta^*$ on M_2. Now from the way μ is defined, it is clear that $\mu(x) = \mu(Tx)$ and therefore that $\mu E^i = \mu TE^i$. Consequently

$$\Delta^* = \mu(\Sigma(E^i + \cdots + T^{q-1}E^i) = q\mu(\Sigma E^i) = 0 \bmod q$$

so that $\Delta \sim 0 \bmod q$, which is impossible. This completes the proof for the case in which T preserves orientation; minor modifications yield a proof for the orientation-reversing case.

2. Suppose that the transformation T of period q operating in M_2 admits at least one fixed point, say x_0. Then for points x near x_0, the functions $u(x)$ and $\omega(x)$ introduced above are uniquely defined and $\omega(x)$ is continuous. Let x vary continuously along a path joining x_0 to a point say x_1 different from x_0. Then $\omega(x)$ will remain defined and continuous so long as the orbit of x is contained in a cap (the cap depending of course on x). As a consequence of the theorem above, $\omega(x)$ is not defined for all x. Suppose x_1 is a point for which $\omega(x)$ is not defined. Then as x moves along the path x_0x_1, we must come to a last point $\bar{x}$ whose orbit is contained in a cap. From elementary continuity consideration, however, it is fairly clear that the orbit of $\bar{x}$ will be contained in a hemisphere H and that at least two points of the orbit of $\bar{x}$ lie on the great

circle boundary of H; if there are exactly two, they must be diametrically opposite each other.

Suppose again that T is a transformation of period q operating in M_2. It can be shown without great difficulty that if the geodesic distance $d(x, Tx)$ is uniformly smaller than $2\pi r/q$, r being the radius of M_2, then each orbit defined by T would be contained in a cap. Since this is impossible, we conclude that there exists at least one point x such that $d(x, Tx) \geqq 2\pi r/q$. In particular, if T admits fixed points, there exists at least one x such that $d(x, Tx) = 2\pi r/q$.

Although these results can be generalized to higher dimensions, a number of technical details enter the proof due to the fact that for spaces of higher dimensions, there is no known method of constructing invariant triangulations. It will be seen in what follows how this difficulty can be met. We shall not, however, pay further attention to transformations of spheres but merely state the results:

Let M_n be an n-sphere of radius r in euclidean $(n + 1)$-space. Every periodic transformation operating in M_n admits at least one orbit which is not contained in any cap (a cap being any closed region on M_n bounded by an $(n - 1)$-sphere which is smaller than a great $(n - 1)$-sphere). *Suppose T is a transformation of period q operating in M_n. If T admits fixed points, there exists at least one point x such that $d(x, Tx) = 2\pi r/q$ where d denotes geodesic distance. Moreover, there will exist at least one point y and integer $s(2 \leqq s \leqq q)$ such that s points of the orbit of y lie on some great $(s - 2)$-sphere of M* (a great 0-sphere being a pair of diametrically opposite points).[2]

More General Spaces

3. From now on the word space will mean a Hausdorff space in which every open set is the sum of a countable family of closed sets (Cf. [3], p. 134).

We shall consistently use the letter M to denote a space, T a transformation operating in M, N a set of points in M and $\mathfrak{A}$ a covering of M (i.e. a finite covering by open sets). We shall write "$T < \mathfrak{A}$ over N" if to each point x of N there can be associated a set of $A_x \in \mathfrak{A}$ such that $x + Tx \subset A_x$. We shall write "$T \ll \mathfrak{A}$ over N" if to each x in N there can be associated an $A_x \in \mathfrak{A}$ which contains the orbit of x.

Lemma 1. Suppose M is locally bicompact, N bounded.[3] For a given integer q and covering $\mathfrak{A}$ there exists a covering $\mathfrak{A}'$ such that every periodic T of period q which satisfies the relation $T < \mathfrak{A}'$ over N satisfies also the relation $T \ll \mathfrak{A}$ over N.

Lemma 2. Suppose M is locally bicompact and H, K are bounded open

[2] The methods suggested by the example we have considered can also be made to yield the theorem that the orbits of a periodic transformation operating in euclidean n-space cannot be uniformly bounded.

[3] A set will be called bounded if its closure is bicompact.

sets such that $\bar{K} \subset H$. There exists a covering $\mathfrak{A}$ such that if $T \ll \mathfrak{A}$ over N (T not necessarily periodic) then

$$\sum_n T^n \bar{K} \subset H.$$

The proofs of lemmas 1 and 2 are elementary.

4. Regularity. Cycles and homologies are to be understood in the sense of Čech (see [2]). Let M be a space, $\mathfrak{g}$ a coefficient group for cycles and homologies in M and n an integer $\geqq 0$. We shall say that M possesses property p_n over $\mathfrak{g}$ at a point d if there exists a neighborhood $D(d)$ and an n-cycle Ω mod $M - D$ with coefficients in $\mathfrak{g}$ such that (1) for no neighborhood $D'(d)$ contained in D is $\Omega \sim 0$ mod $M - D'$ (over $\mathfrak{g}$); (2) for every neighborhood $B(d) \subset D$, there exists a neighborhood $B'(d) \subset B$ such that over $\mathfrak{g}$, Ω is a basis for n-cycles mod $M - B$ relative to homologies mod $M - B'$ while (3) cycles mod $M - B$ of dimension exceeding n are ~ 0 mod $M - B'$. A neighborhood $D(d)$ and relative cycle Ω satisfying (1) (2) (3) will be said to constitute an n-dimensional *fundamental pair* (Ω, D) for d over $\mathfrak{g}$. Clearly if (Ω, D) is an n-dimensional fundamental pair for d over $\mathfrak{g}$, then so is the pair obtained by replacing D by a smaller neighborhood of d.

We shall say that M possesses property q_n over $\mathfrak{g}$ at d if for every neighborhood $A(d)$ there exists a neighborhood $A'(d) \subset A$ such that if C is a non-empty open subset of A', every n-cycle mod $M - A$ in $M - C$ is ~ 0 mod $M - A'$ (all coefficients being in $\mathfrak{g}$).

It will be convenient to call M *n-regular* over $\mathfrak{g}$, if it possesses properties p_n and q_n at each of its points.[4]

We can now state Newman's theorem for spaces of a very general type.

THEOREM I. *Let M be a connected locally bicompact finite dimensional space, N a bounded open set in M and q an integer $\geqq 2$. Let p be a prime factor of q and $\mathfrak{g}_p$ the additive group of integers reduced modulo p. If M is n-regular over $\mathfrak{g}_p$, there exists a covering $\mathfrak{A}$ of M such that no periodic transformation T of period q operating in M can satisfy the relation $T < \mathfrak{A}$ over N.*

The covering $\mathfrak{A}$ necessarily depends on q as one can see from simple examples. But if one replaces the relation $T < \mathfrak{A}$ by the relation $T \ll \mathfrak{A}$ and sharpens the hypothesis on M by requiring that regularity possess a certain uniformity relative $\mathfrak{g}_2, \mathfrak{g}_3, \mathfrak{g}_5, \cdots$ then there can be chosen an $\mathfrak{A}$ independent of q. We shall make this more precise. Suppose that there exist a neighborhood D of d such that for every prime p there exist an n-dimensional fundamental pair (Ω^p, D) over $\mathfrak{g}_p$ for d. We shall say that the fundamental pairs (Ω^p, D), $p = 2, 3, 5, \cdots$, constitute a *fundamental family $F(D, d, n)$*.

Let $F(D, d, n)$ be a fundamental family of fundamental pairs (Ω^p, D). For each p, conditions (1) (2) (3) in the definition of fundamental pair, are satisfied

---

[4] Properties p_n and q_n are slightly weaker than P_n, Q_n in our paper [4].

relative to $\mathfrak{g}_p$. We shall say that conditions (2) and (3) are satisfied *uniformly* by $F(D, d, n)$ if B' is independent of B,—that is, if for every $B(d)$ there is a $B'(d) \subset B$, B' independent of p, such that for every p, Ω^p is a basis over $\mathfrak{g}_p$ for n-cycles mod $M - B$ relative to homologies mod $M - B'$ while cycles mod $M - B$ over $\mathfrak{g}_p$ of dimensions exceeding n are 0 mod $M - B'$. Consider a neighborhood $D'(d) \subset D$. Then for each p, $\Omega^p \not\sim 0 \bmod M - D'$ over $\mathfrak{g}_p$ (condition (1)). Hence for each p there exists a covering $\mathfrak{U}$ of M such that

$$\text{(A)} \qquad \Omega^p(\mathfrak{U}) \not\sim 0 \bmod M - D' \text{ over } \mathfrak{g}_p .$$

If for a given D' there exists a $\mathfrak{U}$ independent of p such that (A) is satisfied for each p, we shall say that $F(D, d, n)$ satisfies condition (1) uniformly. A fundamental family which satisfies (1) (2) and (3) uniformly will be called a *uniform family*. It is easy to see that if $F(D, d, n)$ is a uniform family of fundamental pairs (Ω^p, D), then the fundamental pairs obtained by replacing D by a smaller neighborhood D_1 of d constitute a uniform family $F(D_1, d, n)$.

If for every point d of M there exists a uniform family $F(D, d, n)$, $D = D(d)$, we shall say that property p_n is satisfied uniformly in M. Similarly, property q_n is satisfied uniformly if for each point d and neighborhood $A(d)$, there exists an $A'(d)$ such that if C is a non-null open subset of A and p an arbitrary prime, n-cycles mod $M - A$ in $M - C$ with coefficients in $\mathfrak{g}_p$ are ~ 0, mod $M - A$ (over $\mathfrak{g}_p$). We shall say that M is uniformly n-regular with respect to arbitrary moduli if properties p_n and q_n are satisfied uniformly in M. It can be shown without great difficulty that locally euclidean n-dimensional spaces are uniformly n-regular.

THEOREM II. *Let M be a locally bicompact finite dimensional space and N a bounded open set in M. If M is uniformly n-regular with respect to arbitrary moduli, there exists a covering $\mathfrak{A}$ of M such that no periodic transformation T operating in M can satisfy the relation $T \ll \mathfrak{A}$ over N.*

With regard to theorem I, it is sufficient to prove it for the case in which $q = p =$ a prime. For suppose that $q = hp$, p a prime, and that M is n-regular over $\mathfrak{g}_p$. Then if T is of period q, T^h is of period p and therefore, if $\mathfrak{A}$ is suitably chosen, T can not satisfy the relation $T^h < \mathfrak{A}$ over N. By lemma 1 there exists a covering $\mathfrak{A}'$ such that if $T < \mathfrak{A}'$ over N, then $T \ll \mathfrak{A}$ over N. Since the second of these relations would imply that $T^h < \mathfrak{A}$ over N, T can not satisfy the first. Even simpler considerations make it clear that in order to prove theorem II we need only show that an $\mathfrak{A}$ exists such that no transformation of *prime* period can satisfy the relation $T \ll \mathfrak{A}$ over N.

We remark further that were it not for the fact that in theorem I we assume regularity relative to a single group $\mathfrak{g}_p$, theorem I would be a consequence of theorem II and lemma 1. But it will be seen in the proof of theorem II that in proving the non-existence of small transformations of any *given* prime period p, only the hypothesis of regularity over $\mathfrak{g}_p$ is used. Thus the proof of theorem I is contained essentially in that of theorem II and will therefore not be given separately.

5. It will simplify matters considerably if we carry out the proof of theorem II under the assumption that M is bicompact and then indicate the modifications which will extend the proof to the locally bicompact case. Let us assume then that M is a bicompact space satisfying the hypotheses of theorem II, and that N is an open set in M. Let $m = \dim M$; evidently[5] $m \geqq n$.

Let T be a periodic transformation of prime period p operating in M. We shall denote by L^T the totality of points of M which are left fixed by T, and by $\{\mathfrak{U}^T\}$ the totality of *special* coverings ([3], page 136) of M relative to T. For our present purposes, the essential properties of a special covering $\mathfrak{U}^T$ are the following: (1) the component sets of $\mathfrak{U}^T$ are permuted among themselves by T so that T induces a simplicial transformation of the *complex* (nerve of) $\mathfrak{U}^T$ into itself; (2) the simplexes of $\mathfrak{U}^T$ which are invariant under T form a subcomplex $\mathfrak{U}_0^T$; (3) the simplexes of $\mathfrak{U}_0^T$, and these only, are in[6] L^T; (4) each simplex in $\mathfrak{U}^T - \mathfrak{U}_0^T$ is of dimension $\leqq m$. It is evident that a necessary and sufficient condition that a $\mathfrak{U}^T$-chain be in L^T (in the sense of footnote 5) is that it be in the subcomplex $\mathfrak{U}_0^T$. Moreover, every chain in $\mathfrak{U}^T - \mathfrak{U}_0^T$ is in $M - L^T$, but the converse is not necessarily true. The totality of coverings $\mathfrak{U}^T$ is a complete system ([3], p. 137) and therefore can be used for defining all homology relations in M.

A transformation T of prime period p operating in M induces a boundary-preserving operation (also denoted by T) on the $\mathfrak{U}^T$-chains in M. Let δ, σ denote the operators $1 - T$ and $1 + T + \cdots + T^{p-1}$. We shall use the symbols ρ, $\bar{\rho}$ to stand either for δ, σ or σ, δ respectively. When the coefficient group $\mathfrak{g}$ is $\mathfrak{g}_p$, all $\mathfrak{U}^T$-chains in $\mathfrak{U}_0^T$ are annihilated by ρ and $\bar{\rho}$. For if E is a $\mathfrak{U}_0^T$-simplex, then $\delta E = E - E = 0$ and $\sigma E = pE = 0 \pmod p$. From this remark follows

LEMMA 3. If $\mathfrak{g} = \mathfrak{g}_p$ where p is the period of T, assumed to be prime, then every $\mathfrak{U}^T$-chain of the form ρX (or $\bar{\rho} X$) is in $\mathfrak{U}^T - \mathfrak{U}_0^T$, hence in $M - L$.

Suppose now that B is an invariant set, that is, a set which coincides with its image under T. Suppose there exists a $\mathfrak{U}^T$-chain of the form ρX which is a cycle mod B. If ρX is, modulo B, the boundary of a $\mathfrak{U}^T$-chain of the same form, we shall write $\rho X \simeq 0 \bmod B$. In particular, if $\rho X = 0 \bmod B$, then $\rho X \simeq 0 \bmod B$.

---

[5] The regularity hypotheses on M imply that for each p the modulo p dimension (Alexandroff) of M is n. This however, so far as we know, does not imply that $\dim M = n$. If it were definitely assumed that $\dim M$ equals n (not merely that it is finite), the proof of Theorem II would be somewhat shorter since the existence of the relative cycles in propositions A), C), and F) could be established without the use of lemma 4. We prefer to consider the more general situation on account of possible connections with the problem mentioned in the introduction. For in this problem one is led to consider transformations in spaces about whose dimension nothing is as yet known.

[6] A $\mathfrak{U}$-simplex $(U_0 U_1 \cdots U_h)$ is in the set A if $U_0 U_1 \cdots U_h A \neq 0$. A $\mathfrak{U}$-chain is in A if each of its simplexes is in A. A cycle is in A if each of its coordinates is in A.

LEMMA 4. Let $\mathfrak{g} = \mathfrak{g}_p$ and suppose that ρX_h, $\bar{\rho}X_{h-1}$ are $\mathfrak{U}^T$-cycles mod B such that

$$\beta X_h = \bar{\rho}X_{h-1} \bmod B$$

(β being the boundary operator). If $\rho X_h \simeq 0 \bmod B$, then $\bar{\rho}X_{h-1} \simeq 0 \bmod B$.

Proof. Suppose in fact that there exists a relation of the form $\beta\rho X_{h+1} = \rho X_h \bmod B$. Write

$$\beta X_{h+1} = X_h - Z. \tag{1}$$

On operating on both sides of (1) by ρ we find that $\rho Z = 0 \bmod B$, hence by [3], section 3.11, we may write $Z = \bar{\rho}W + W^L \bmod B$ with $W^L \subset L$. Therefore, if we operate on both sides of (1) by β we obtain

$$0 = P + Q \bmod B, \qquad P = \bar{\rho}(X_{h-1} - \beta W), \qquad Q = \beta W^L.$$

Since Q is in $\mathfrak{U}_0^T$ and since, by lemma 3, P is in $\mathfrak{U}^T - \mathfrak{U}_0^T$, the chains P and Q have no common simplex, hence $P = Q = 0 \bmod B$. In particular $\beta\bar{\rho}W = \bar{\rho}X_{h-1} \bmod B$ so that $\bar{\rho}X_{h-1} \simeq 0 \bmod B$.

6. We now prove[7] theorem II for the bicompact case by establishing a sequence of propositions concerning chains and cycles associated with transformations of prime period. Let M be a connected bicompact space, uniformly n-regular relative to arbitrary moduli. Let N be an open subset of M and let $m = \dim M$. Let it be understood in propositions A) B) C) D) that T is an arbitrary but fixed transformation of prime period p operating in M, and that the coefficient group is $\mathfrak{g}_p$.

A) Let d be a point of M. Suppose there exists for d an n-dimensional fundamental pair (Ω, G) such that G is invariant under T. Let $k = m - n + 1$ and suppose further that there exist invariant neighborhoods of d:

$$G \supset A_0 \supset A_1 \supset \cdots \supset A_k$$

such that (1) Ω is a basis for n-cycles mod $M - G$ relative to homologies mod $M - A_0$; (2) for $i = 0, \cdots, k$, cycles mod $M - A_{i-1}$ of dimension $> n$ are ~ 0 mod $M - A_i$. Then there exists an n-cycle Δ mod $M - A_k$ of the form

$$\{\sigma\Gamma(\mathfrak{U}^T) + \Gamma^L(\mathfrak{U}^T)\}, \qquad \Gamma^L(\mathfrak{U}^T) \subset L^T$$

such that $\Delta \sim \Omega \bmod M - A_k$.

Proof. Since G is invariant, $T\Omega$ is a cycle mod $M - G$ and hence $T\Omega \sim x\Omega$ mod $M - A_0$ where $x \in \mathfrak{g}_p$. Since A_0 is also invariant, this homology holds if $T\Omega$ is substituted for Ω. On repeated substitution, we obtain: $\Omega = T^p\Omega \sim x^p\Omega \bmod M - A_0$, hence $x^p = 1$, $x = 1$, so that

$$\delta\Omega \sim 0 \bmod M - A_0. \tag{1}$$

[7] The proof given here is somewhat shorter and simpler than an earlier proof proposed by the author, thanks to valuable suggestions from Professor Norman Steenrod.

Assume for the moment that $m > n$. Let $\mathfrak{U}^T$ be a special covering and i an integer in the range $1, \cdots, k$ and h an integer in the range $n + 1, \cdots, m$. From the Čech homology theory, there exists a special refinement $\mathfrak{V}^T$ of $\mathfrak{U}^T$ such that the projection into $\mathfrak{U}^T$ of any h-dimensional $\mathfrak{V}^T$-cycle mod $M - A_{i-1}$ will be the $\mathfrak{U}^T$-coordinate of a Čech cycle mod $M - A_{i-1}$, hence will (by the definition of the A's) be ~ 0 mod $M - A_i$ (see [2]). It is obvious that $\mathfrak{V}^T$ can be chosen to be independent of h since the range of h is finite.—Now let $\mathfrak{U}^T$ be an arbitrarily chosen special covering. From what has been said, we may choose special coverings

$$\mathfrak{U}^T = \mathfrak{U}_k \supset \mathfrak{U}_{k-1} \supset \cdots \supset \mathfrak{U}_0$$

such that for $i = 1, \cdots, k$, the projection into $\mathfrak{V}_i$ of a $\mathfrak{U}_{i-1}$-cycle mod $M - A_{i-1}$ of dimension exceeding n but not m, is ~ 0 mod $M - A_i$. Let π_i be a projection $\mathfrak{U}_{i-1} \rightarrow \mathfrak{U}_i$ which is permutable with T,—it is easy to see that such projections exist (see [3], p. 138). Let $\rho_0, \rho_1, \cdots$ stand alternately for δ, σ beginning with $\rho_0 = \delta$.

Putting $\Omega = \Omega_n$, it follows from (1) that $\rho_0\Omega_n(\mathfrak{U}_0) \sim 0$ mod $M - A_0$. Hence $\pi_1\rho_0\Omega_n(\mathfrak{U}_0) \sim 0$ mod $M - A_0$, say (with $i = 0$)

$$\beta\Omega_{n+i+1}(\mathfrak{U}_{i+1}) = \pi_{i+1}\rho_i\Omega_{n+i}(\mathfrak{U}_i) \bmod M - A_i. \tag{2}$$

Now $\rho_1\Omega_{n+1}(\mathfrak{U}_1)$ is a cycle mod $M - A_0$ since its boundary mod $M - A_0$ is $\rho_0\rho_1(\pi_1\Omega_n) = \delta\sigma\pi_1\Omega_n$ which vanishes since $\delta\sigma = (1 - T)(1 + \cdots + T^{p-1}) = 1 - T^p = 0$. From the definition of the $\mathfrak{U}$'s we have $\pi_2\rho_1\Omega_{n+1}(\mathfrak{U}_1) \sim 0$ mod $M - A_1$. Hence there exists a relation (2) with $i = 1$. On repeating the construction we obtain relations (2) corresponding to $i = 0, \cdots, k$. Let

$$\begin{aligned} Y_{n+i} &= \pi_k\pi_{k-1} \cdots \pi_{i+1}\Omega_{n+i}(\mathfrak{U}_i) \qquad (i = 0, \cdots, k - 1) \\ Y_{n+k} &= \Omega_{n+k}(\mathfrak{U}_k). \end{aligned} \tag{3}$$

Then from (2) we have

$$\beta Y_{n+i} = \rho_{i-1}Y_{n+i-1} \bmod M - A_i \qquad (i = 1, \cdots, k). \tag{4}$$

Now ρ_iY_{n+i} is a $\mathfrak{U}_k$-cycle mod $M - A_i$, hence mod $M - A_k$. The particular cycle ρ_kY_{n+k} is null. For by lemma 3, this cycle lies in $\mathfrak{U}_k - (\mathfrak{U}_k)_0$. But this is impossible unless $\rho_kY_{n+k} = 0$ since $n + k = m + 1$ whereas by the properties of special coverings, the simplexes of $\mathfrak{U}_k - (\mathfrak{U}_k)_0$ are of dimension $\leqq n$. We conclude therefore that $\rho_kY_{n+k} \simeq 0$ mod $M - A_k$. Hence by successive applications of lemma 4, the relations (4) yield

$$\rho_iY_{n+i} \simeq 0 \bmod M - A_k \qquad (i = k, \cdots, 0).$$

In particular $\rho_0Y_n \simeq 0$ mod $M - A_k$ and hence there exists a relation of the form

$$\beta\rho_0X_{n+1}(\mathfrak{U}_k) = \rho_0Y_n(\mathfrak{U}_k) \bmod M - A_k.$$

Write

$$\beta X_{n+1}(\mathfrak{U}_k) = Y_n(\mathfrak{U}_k) - H(\mathfrak{U}_k). \tag{5}$$

On operating on both sides of (5) by ρ_0, we find that $\rho_0 H = \delta H = 0$ mod $M - A_k$. Hence by [3], section 3.11, there exist chains $\Gamma(\mathfrak{U}_k)$, $\Gamma^L(\mathfrak{U}_k)$, the latter in L, such that

$$H(\mathfrak{U}_k) = \sigma\Gamma(\mathfrak{U}_k) + \Gamma^L(\mathfrak{U}_k) \bmod M - A_k. \tag{6}$$

From (3) with $i = 0$, we have $Y_n \sim \Omega_n$ mod $M - A_k$. Consequently from (5) and (6) we have

$$\Omega_n(\mathfrak{U}_k) \sim \sigma\Gamma(\mathfrak{U}_k) + \Gamma^L(\mathfrak{U}_k) \bmod M - A_k$$

which, since $\mathfrak{U}_k$ is an arbitrary special covering, establishes A) under the assumption that $m > n$. The proof for the case $m = n$ is not essentially different and we therefore omit further details.

B) In the cycle Δ of proposition A) the chains $\sigma\Gamma(\mathfrak{U}^T)$ and $\Gamma^L(\mathfrak{U}^T)$ are $\mathfrak{U}^T$-cycles mod $M - A_k$.

For in any case $\beta\Delta$ equals

$$\sigma\beta\Gamma(\mathfrak{U}^T) + \beta\Gamma^L(\mathfrak{U}^T) = 0 \bmod M - A_k. \tag{1}$$

Now $\beta\Gamma^L(\mathfrak{U}^T) \subset \mathfrak{U}_0^T$ whereas, by lemma 3, $\sigma\beta\Gamma(\mathfrak{U}^T) \subset \mathfrak{U}^T - \mathfrak{U}_0^T$. Hence the two chains in the left member of (1) have no common simplex and therefore vanish separately mod $M - A_k$.

C) Let d be a point in L^T. There exists for d over $\mathfrak{g}_p$ an n-dimensional fundamental pair (Δ, A) such that $A(d)$ is invariant and Δ is of the form

$$\{\sigma\Gamma(\mathfrak{U}^T) + \Gamma^L(\mathfrak{U}^T)\}, \qquad \Gamma^L(\mathfrak{U}^T) \subset L^T.$$

For by the hypothesis of uniform n-regularity there exists for d over $\mathfrak{g}_p$ an n-dimensional fundamental pair (Ω, D). Now since d is a point of L^T, any given neighborhood $C(d)$ will contain an invariant neighborhood of d. Such a neighborhood, for example, is the intersection of $C, TC, \cdots, T^{p-1}C$. Thus D can be replaced by an invariant neighborhood of d. Moreover, from the n-regularity of M over $\mathfrak{g}_p$, there exist neighborhoods $A_0, \cdots, A_k$ of d satisfying the conditions stated in A). The set $A = A_k$ forms together with the cycle Δ of A) the desired fundamental pair.

D) L^T is nowhere dense in M.

For, let $L = L^T$ and let L_0 be the maximal open subset of L. Assume, contrary to D) that $L_0 \neq 0$. Then $\bar{L}_0 - L \neq 0$, otherwise L_0 would be both open and closed, hence identical with M so that T would be the identity. Let d be a point in $\bar{L}_0 - L_0$ and let (Δ, A) be the fundamental pair for d of proposition C). On account of the n-regularity of M over $\mathfrak{g}_p$ there exists a neighborhood $A'(d) \subset A$ such that if C is a non-empty subset of A, n-cycles mod $M - A$ in $M - C$ are ~ 0 mod $M - A'$. Suppose in particular that $C = A' - A'L$. Then by the choice of d, $C \neq 0$. Moreover $L \subset M - C$ since $L(A' - A'L) = 0$.

Therefore $\Gamma^L(\mathfrak{U}^T) \subset M - C$ for every $\mathfrak{U}^T$. Let $\mathfrak{U}$ be an arbitrary but fixed $\mathfrak{U}^T$. If $\mathfrak{B} \in \{\mathfrak{U}^T\}$ is a suitably chosen refinement of $\mathfrak{U}^T$ and π an arbitrarily chosen projection $\mathfrak{B} \to \mathfrak{U}$, then $\pi\Gamma^L(\mathfrak{B})$ is the $\mathfrak{U}$-coordinate of a Čech cycle mod $M - A$ in $M - C$ (see proof of A)), hence is ~ 0 mod $M - A'$. (That $\Gamma^L(\mathfrak{B})$ is a cycle mod $M - A$ follows from proposition C)). Next take for C the non-empty set L_0A'. Then $\sigma\Gamma \subset M - C$ since in any case $\sigma\Gamma \subset M - L$ by lemma 3 and $M - L \subset M - L_0A' = M - C$. Hence as above, for a suitable special refinement $\mathfrak{B}_1$ of $\mathfrak{U}$ and projection $\pi_1 : \mathfrak{B}_1 \to \mathfrak{U}$, we have $\pi_1\sigma\Gamma(\mathfrak{B}_1) \sim 0$ mod $M - A'$. We may assume that $\mathfrak{B} = \mathfrak{B}_1$ for both coverings can be replaced by a common (special) refinement. If we then take $\pi = \pi_1$ we have

$$\pi(\sigma\Gamma(\mathfrak{B}) + \Gamma^L(\mathfrak{B})) \sim 0 \bmod M - A'$$

and hence the $\mathfrak{U}$-coordinate of Δ is ~ 0 mod $M - A'$. Since $\mathfrak{U}$ is an arbitrary special covering, this implies that $\Delta \sim 0$ mod $M - A'$ which is impossible.

E) Let d be a point in N and let $F(D, d, n)$ be a uniform family of n-dimensional fundamental pairs (Ω^q, D) $(q = 2, 3, 5, \cdots)$ such that $D(d) \subset N$. There exists a neighborhood $B(d) \subset D$ and a covering $\mathfrak{B}$ of M such that if T is a periodic transformation of prime period p satisfying the relation: $T \ll \mathfrak{B}$ over N, there exist chains of the form $\sigma\Gamma(\mathfrak{U}^T)$ with coefficients in $\mathfrak{g}_p$ such that

$$\{\sigma\Gamma(\mathfrak{U}^T)\} \sim \Omega^p \bmod M - B.$$

To prove this, choose a neighborhood $D'(d) \subset D$ such that for each prime q, Ω^q is a basis for n-cycles mod $M - D$ relative to homologies mod $M - D'$ (coefficients in $\mathfrak{g}_q$). On replacing D' by a smaller neighborhood of d if necessary, we may assume that $\bar{D}' \subset D$. Next, choose a neighborhood $D''(d)$ related to D' in the same way that D' is related to D. Now let $k = m - n + 1$ and choose neighborhoods $B_i(d)$, $B_i'(d)$ $(i = 0, \cdots, k)$ contained in D such that

$$B_i' \supset B_{i+1} \qquad (i = 0, \cdots, k-1)$$

$$B_i \supset \bar{B}_i' \qquad (i = 0, \cdots, k)$$

and such that, for each q, cycles mod $M - B_i'$ over $\mathfrak{g}_q$ of dimension $> n$ are ~ 0 mod $M - B_{i+1}$ $(i = 0, \cdots, k - 1)$. By repeated applications of lemma 2 there exists a covering $\mathfrak{B}$ such that if $T \ll \mathfrak{B}$ over N, then $D \supset \sigma D'$ and $B_i \supset \sigma B_i'$, so that

$$D \supset \sigma D' \supset D' \supset D'' \supset B_0 \supset \sigma B_0' \supset B_0' \supset B_1 \supset \sigma B_1' \supset B_1' \supset \cdots.$$

In particular, if we write $G = \sigma D'$ and $A_i = \sigma B_i'$, we will have

$$D \supset G \supset A_0 \supset A_1 \supset \cdots \tag{1}$$

$$G \supset D' \supset D'' \supset A_0. \tag{2}$$

The sets $G, A_0, A_1, \cdots$ depend on T since σ does but the sets D', D'', B_i, B_i' do not. Consider $B_k(d)$. We can choose a neighborhood $B(d) \subset B_k$ such that

if C is a non-empty open subset of B and q a prime, then over $\mathfrak{g}_q$, n-cycles mod $M - B$ in $M - C$ are ~ 0 mod $M - B'$.

Now consider a definite T of prime period p such that $T \ll \mathfrak{B}$ over N. Since $G \subset D$, (Ω^p, G) is a fundamental pair over $\mathfrak{g}_p$. Moreover, it follows from (2) and the relation between D' and D'' that Ω^p is a basis over $\mathfrak{g}_p$ for n-cycles mod $M - G$ relative to homologies mod $M - A_0$. Finally, cycles mod $M - A_i$ over $\mathfrak{g}_p$ of dimension $> n$ are ~ 0 mod $M - A_{i+1}$; for, such cycles are cycles mod $M - B_i'$ since $B_i' \subset \sigma B_i' = A_i$ and as such, they are ~ 0 mod $M - B_{i+1}$, hence mod $M - A_{i+1}$ since $A_{i+1} = \sigma B_{i+1}' \subset B_{i+1}$. From these facts and the relations (1) it follows that the sets $G, A_0, \cdots, A_k$ satisfy the conditions in proposition A) and consequently there exists a cycle Δ mod $M - A_k$ over $\mathfrak{g}_p$ of the form

$$\{\sigma\Gamma(\mathfrak{U}^T) + \Gamma^L(\mathfrak{U}^T)\} \tag{3}$$

such that $\Delta \sim \Omega^p$ mod $M - A_k$.

Let γ^T denote an arbitrary n-cycle mod $M - A_k$ in L^T, coefficients in $\mathfrak{g}_p$. γ^T is a cycle mod $M - B_k'$ since $B_k' \subset \sigma B_k' = A_k$. The open set $B_k' - L^T B_k'$ is non-empty since L^T is nowhere dense (proposition D)). Hence γ^T, being in L^T, hence in $M - (B_k' - L^T B_k')$, is ~ 0 mod $M - B$ by choice of B. In short, n-cycles mod $M - A_k$ (over $\mathfrak{g}_p$) in L^T are ~ 0 mod $M - B$. Now for each $\mathfrak{U}^T$, $\Gamma^L(\mathfrak{U}^T)$ is a cycle mod $M - A_k$ in L^T (proposition B)). If $\mathfrak{B}^T$ is a refinement of $\mathfrak{U}^T$ and π a projection $\mathfrak{B}^T \to \mathfrak{U}^T$, the $\mathfrak{U}^T$-coordinate of Δ can be replaced by the projection of its $\mathfrak{B}^T$-coordinate by π. But if $\mathfrak{B}^T$ is suitably chosen, $\pi\Gamma^L(\mathfrak{B}^T)$ will be the $\mathfrak{U}^T$-coordinate of a Čech cycle mod $M - A_k$ in L^T, coefficients in $\mathfrak{g}_p$, hence will be ~ 0 mod $M - B$. Hence we may assume in (3) that $\{\Gamma^L(\mathfrak{U}^T)\} \sim 0$ mod $M - B$. Hence $\Omega^p \sim \{\sigma\Gamma(\mathfrak{U}^T)\}$ mod $M - B$. This establishes E).

F) Let $\mathfrak{U}$ be a covering of M and let T be a periodic transformation such that $T \ll \mathfrak{U}$ over N. Let N^* be a non-empty open set such that $\bar{N}^* \subset N$. There exists a special refinement $\mathfrak{U}^T$ of $\mathfrak{U}$ and projection π: $\mathfrak{U}^T \to \mathfrak{U}$ such that $\pi T = \pi$ for all $\mathfrak{U}^T$-simplexes in N^*.

The component sets of a given $\mathfrak{U}^T$ fall into "cyclic families," the sets in a given family being images of each other under powers of T. Each family contains exactly p sets or a single (invariant) set. Let cU^T denote the cyclic family to which $U^T(\epsilon\, \mathfrak{U}^T)$ belongs. Let $\nu(\mathfrak{U}^T)$ denote the totality of cyclic families cU^T such that at least one member of cU^T meets N^*. To establish F) it will be sufficient to show that there exists a refinement $\mathfrak{U}^T$ of $\mathfrak{U}$ such that to each cyclic family cU^T in $\nu(\mathfrak{U}^T)$ there can be associated a set $U = U(cU^T)$ of $\mathfrak{U}$, containing all the sets in cU^T. For if π is a projection $\mathfrak{U}^T \to \mathfrak{U}$ carrying each set in cU^T (cU^T a member of $\nu(\mathfrak{U}^T)$) into $U(cU^T)$ and arbitrarily defined for the remaining sets of $\mathfrak{U}^T$, it is clear that $\pi T = \pi$ for $\mathfrak{U}^T$-complexes in N^*. To show that $\mathfrak{U}^T$ exists, let there be associated to each point x in N a set U_x of $\mathfrak{U}$ containing the orbit of x. This is possible since $T \ll \mathfrak{U}$ over N. Evidently we may associate to each x in N a neighborhood $V(x)$ such that the images of $V(x)$ under powers of T are contained in U_x. Let $V(x_1), \cdots, V(x_n)$ be a finite number of

these sets forming a covering of $\bar{N}^*$ and let Y be an open set such that $\bar{N}^* \subset Y$, $\bar{Y} \subset \Sigma V(x_i)$. Let $\mathfrak{X}$ be the totality of sets $V(x_i)$ together with the single set $M - \bar{Y}$. Evidently $\mathfrak{X}$ is a covering of M. Now the closed sets $\bar{N}^*$ and $M - Y$ are disjoint. Hence there can be chosen a $\mathfrak{U}^T$, refinement of $\mathfrak{X}$, such that sets of $\mathfrak{U}^T$ which meet $\bar{N}^*$ will not meet $M - Y$. Consider a cyclic family say cU_1^T in $\nu(\mathfrak{U}^T)$. At least one member of cU^T meets N^*. That member, however, can not meet $M - \bar{Y}$, hence it is contained in one of the sets $V(x_i)$ say $V(x_1)$. Hence its images—the members of cU_1^T—are all contained in U_{x_1}. If we define the function $U(cU)$ to be U_{x_1} for the cyclic family cU_1^T and similarly for the remaining cyclic families of $\nu(\mathfrak{U}^T)$, it will possess the desired properties.

G) There exists a covering $\mathfrak{A}$ of M such that no periodic T of prime period can satisfy the relation $T \ll \mathfrak{A}$ over N.

For let d be a point of N and let $F(D, d, n)$ be a uniform family of n-dimensional fundamental pairs (Ω^p, D). On replacing D by a smaller neighborhood of d if necessary, we may assume that $D \subset N$. Let $B(d)$, $\mathfrak{B}$ be the neighborhood and covering of proposition E) and let $H(d)$ be a neighborhood of d such that $\bar{H} \subset B$. Since $F(D, d, n)$ uniformly satisfies condition (1) in the definition of fundamental pair (section 4) there exists a covering $\mathfrak{U}$ such that for arbitrary prime p,

$$\Omega^p(\mathfrak{U}) \sim 0 \bmod M - H. \tag{1}$$

Let $\mathfrak{B}$ be a covering such that if T is a periodic transformation with $T \ll \mathfrak{B}$ over N, then $\sigma\bar{H} \subset B$ (lemma 2). Let $\mathfrak{A}$ be a common refinement of $\mathfrak{B}$, $\mathfrak{U}$, $\mathfrak{B}$. We shall show that $\mathfrak{A}$ has the required property. Suppose there exists a T of prime period p such that $T \ll \mathfrak{A}$ over N. Let $N^* = \sigma H$. Then $\bar{N}^* \subset B \subset N$. By proposition F) there exists a refinement $\mathfrak{U}^T$ of $\mathfrak{U}$ and projection π: $\mathfrak{U}^T \to \mathfrak{U}$ such that $\pi T = \pi$ for $\mathfrak{U}^T$-simplexes in N^*. Consider the relative cycle $\{\sigma\Gamma(\mathfrak{U}^T)\}$ of proposition E). Let $\Lambda(\mathfrak{U}^T)$ be the subchain of $\Gamma(\mathfrak{U}^T)$ consisting of these simplexes which are in N^*. Then $\Gamma(\mathfrak{U}^T) - \Lambda(\mathfrak{U}^T) \subset M - N^*$. Since N^* is invariant under T, we have also

$$\sigma\Gamma(\mathfrak{U}^T) - \sigma\Lambda(\mathfrak{U}^T) \subset M - N^*.$$

This relation holds if the left member is replaced by its image under π. But since the chain $\sigma\Lambda(\mathfrak{U}^T)$ is in N^*, its image under π is $p\Lambda(\mathfrak{U}^T) = 0 \bmod p$. Hence $\pi\sigma\Gamma(\mathfrak{U}^T) \sim 0 \bmod M - N^*$. Since $\Omega^p(\mathfrak{U}^T) \sim \sigma\Gamma(\mathfrak{U}^T) \bmod M - B$, hence mod $M - H$, we have $\pi\Omega^p(\mathfrak{U}^T) \sim 0 \bmod M - H$. By definition of Čech cycle, $\pi\Omega^p(\mathfrak{U}^T) \sim \Omega^p(\mathfrak{U}) \bmod M - H$ so that $\Omega^p(\mathfrak{U}) \sim 0 \bmod M - H$ which contradicts (1). This concludes the proof of theorem II under the hypothesis that M is bicompact.

7. We conclude with a word concerning the locally bicompact case. In the first place, an examination of the preceding proof shows that in the bicompact case, it would have been sufficient to assume merely that M is uniformly n-regular over N instead of the whole of M. Now suppose that M, N satisfy the

conditions of the theorem but that M is not bicompact. M can be converted into a bicompact space M_1 by the addition of a single point[8] say ∞. In M_1, N has for its counterpart an open set N_1 not containing ∞. If M is uniformly n-regular, then M_1 is at least uniformly n-regular over N_1. Hence there is a covering $\mathfrak{A}_1$ of M_1 such that no periodic transformation T_1 operating in M_1 can satisfy the relation $T_1 \ll \mathfrak{A}_1$ over N_1. Now every periodic T operating in M induces a periodic T_1, of same period, operating in M_1 and leaving ∞ fixed. If $\mathfrak{A}$ is the covering of M obtained from $\mathfrak{A}_1$ by suppressing the point ∞, it is easy to see that the relation: $T \ll \mathfrak{A}$ over N would imply $T \ll \mathfrak{A}_1$ over N_1. Hence $\mathfrak{A}$ is the required covering of M.

The Institute for Advanced Study and Barnard College,
Columbia University

References

(1) M. H. A. Newman, A theorem on periodic transformations of spaces, Quart. Journal 2 (1931) pp. 1–8.
(2) E. Čech, Théorie générale de l'homologie dans un espace quelconque, Fund. Math. 18 (1932) pp. 149–183.
(3) P. A. Smith, Transformations of finite period, Annals of Math. 39 (1938) pp. 127–164.
(4) P. A. Smith, Transformations of finite period. II, Annals of Math. 40 (1939) 690–711.

[8] We may define as neighborhood of ∞ the complement of any bicompact set in M, plus ∞ itself. With this topology M_1, like M, is a connected finite-dimensional space in which open sets are countable sums of closed sets. Moreover, M_1 will be uniformly n-regular except possibly at the point ∞.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

CONTINUOUS MAPPINGS OF INFINITE POLYHEDRA

BY SAMUEL EILENBERG

(Received March 20, 1940)

This paper is concerned with some special cases of the general problem of determining the homotopy classes of continuous mappings of a polyhedron K into a space Y. The well known theorem of H. Hopf solves the problem when K is a finite n-dimensional polyhedron and $Y = S^n$ is the n-sphere. The solution is given in terms of homology. In a recent paper[1] the author has generalized Hopf's result simultaneously in the following three directions:

(a) The hypothesis that K is finite is removed by admitting infinite chains and letting cohomologies replace homologies.

(b) The hypothesis that K is n-dimensional is replaced by a hypothesis concerning the vanishing of certain higher-dimensional cohomology groups.

(c) S^n is replaced by a more general space Y.

Although the use of cohomologies in this problem seems quite natural and justified, the results become more intuitive and easier to apply in the cases when the theorem has a homological interpretation. Such an interpretation is easily available when K is finite but does not seem to be possible generally. Section **2** of this paper gives such an interpretation for a large class of infinite polyhedra. Various applications are given in sections **3–10**. The results of **CM** which are being used here are collected in section **1**.

1. Let K be a finite or infinite cell-complex, i.e. a locally finite polyhedron with a given cell decomposition. K^n will denote the sub-complex of K consisting of all cells of dimension $\leqq n$. Given an abelian group G we shall consider infinite chains in K with coefficients in G and denote the resulting n-dimensional cohomology group by ${}^n\boldsymbol{H}_G(K)$.

Let Y be an arcwise connected topological space and let $y_0 \in Y$. We shall denote by $\pi_i(Y)$ the i^{th} homotopy group of Y with respect to y_0 as origin.[2] We shall assume that $\pi_i(Y) = 0$ for $i < n$. If $n = 1$ we replace this condition by the condition that $\pi_1(Y)$ (= the fundamental group of Y) is abelian.

The set of all continuous mappings $f(K) \subset Y$ will be denoted by Y^K. Given $f_0, f_1 \in Y^K$ we use the notation $f_0 \simeq f_1$ to denote that f_0 and f_1 are homotopic.

(1.1) *Given $f \in Y^K$ there is an $f' \in Y^K$ such that $f \simeq f'$ and $f'(K^{n-1}) = y_0$* **[CM,** p. 241]

---

[1] Cohomology and continuous mappings, these *Annals*, vol. 41 (1940), pp. 231–251. We shall refer to this paper as to **CM**.

[2] This group was introduced by W. Hurewicz, *Proc. Akad. Amsterdam* 38 (1935), p. 113. The definition given in **CM**, p. 235, seems to suit our purposes best.

Given $f \epsilon Y^K$ such that $f(K^{n-1}) = y_0$ we construct an n-dimensional chain $d^n(f)$ in K with coefficients from $\pi_n(Y)$ as follows:

Let σ^n be an oriented n-cell of K. Since the boundary of σ^n is carried by f into the point y_0, the mapping $f(\sigma^n) \subset Y$ determines uniquely an element $d(f, \sigma^n)$ of $\pi_n(Y)$. We define

$$d^n(f) = \sum_i d(f, \sigma_i^n)\sigma_i^n$$

We have proved that

(1.2) *$d^n(f)$ is a cocycle in K.* [**CM**, p. 240, Hom. Th. 1]

(1.3) *If $f_0, f_1 \epsilon Y^K$, $f_0(K^{n-1}) = f_1(K^{n-1}) = y_0$ and $f_0 \simeq f_1$ then $d^n(f_0)$ and $d^n(f_1)$ are cohomologous in K.* [**CM**, p. 242, Hom. Th. 3]

Given an arbitrary $f \epsilon Y^K$ there is by (1.1) a mapping $f' \epsilon Y^K$ such that $f \simeq f'$ and $f'(K^{n-1}) = y_0$. From (1.3) it follows that the cohomology class of the cocycle $d^n(f')$ is independent of the choice of f'. We denote this cohomology class by $\boldsymbol{d}^n(f)$. It follows from (1.3) that

(1.4) *If $f_0, f_1 \epsilon Y^K$ and $f_0 \simeq f_1$ then $\boldsymbol{d}^n(f_0) = \boldsymbol{d}^n(f_1)$.*

It follows that for each homotopy class Φ of Y^K the cohomology class $\boldsymbol{d}^n(\Phi)$ is uniquely defined.

The following theorem includes all the further results of **CM** which will be needed in the sequel [**CM**, p. 243].

THEOREM 1. *Let K be a finite or infinite cell-complex and Y an arcwise connected topological space such that $\pi_i(Y) = 0$ for $i < n$. If $n = 1$ this condition is replaced by the condition that Y is i-simple*[3] *for $i = 1, 2, \cdots, \dim K$.*

(1.5) *If ${}^{i+1}\boldsymbol{H}_{\pi_i(Y)}(K) = 0$ for $i = n + 1, n + 2, \cdots$ then given a cocycle d^n of K with coefficients in $\pi_n(Y)$ there is a mapping $f \epsilon Y^K$ such that $f(K^{n-1}) = y_0$ and $d^n(f) = d^n$.*

(1.6) *If ${}^{i}\boldsymbol{H}_{\pi_i(Y)}(K) = 0$ for $i = n + 1, n + 2, \cdots$ then, given $f_0, f_1 \epsilon Y^K$ such that $f_0(K^{n-1}) = f_1(K^{n-1}) = y_0$, we have $f_0 \simeq f_1$ if and only if $d^n(f_0)$ and $d^n(f_1)$ are cohomologous in K.*

(1.7) *If ${}^{i}\boldsymbol{H}_{\pi_i(Y)}(K) = {}^{i+1}\boldsymbol{H}_{\pi_i(Y)}(K) = 0$ for $i = n + 1, n + 2, \cdots$ then the homotopy classes Φ of Y^K are in a (1-1)-correspondence with the elements of the cohomology group ${}^{n}\boldsymbol{H}_{\pi_n(Y)}(K)$. The correspondence is determined by the operation $\boldsymbol{d}^n(\Phi)$.*

2. Let ${}^n\mathcal{H}^I(K)$ be the n^{th} homology group of K obtained using finite chains, and ${}^n\mathcal{H}^I(Y)$ the n^{th} homology group of Y obtained using finite singular chains. In both cases the group I of integers is taken as the coefficient group.

It is well known that every $f \epsilon Y^K$ induces a homomorphic mapping h_f^n of ${}^n\mathcal{H}^I(K)$ into a subgroup of ${}^n\mathcal{H}^I(Y)$ and also that if $f_0, f_1 \epsilon Y^K$ and $f_0 \simeq f_1$ then $h_{f_0}^n = h_{f_1}^n$. It follows that for every homotopy class Φ of Y^K the homomorphism h_Φ^n is uniquely defined.

---

[3] See **CM**, p. 235. The condition was introduced by the author (*Fund. Math.* 32 (1939), pp. 167-175) and it concerns the mutual behavior of $\pi_1(Y)$ and $\pi_i(Y)$. In particular Y is i-simple if one of these groups vanishes. Y is 1-simple if and only if $\pi_i(Y)$ is abelian.

Let us assume that K is connected, Y is arcwise connected and that $\pi_i(K) = \pi_i(Y) = 0$ for $i < n$. If $n = 1$ the last condition will be replaced by the condition that $\pi_1(K)$ and $\pi_1(Y)$ are abelian. In this case it was proved by W. Hurewicz that the groups $\pi_n(K)$ and ${}^n\mathcal{H}^I(K)$ can be considered identical.[4] Similarly for $\pi_n(Y)$ and ${}^n\mathcal{H}^I(Y)$. The homomorphism h_f^n induced by a mapping $f \epsilon Y^K$ becomes then a homomorphic mapping of $\pi_n(K)$ into a subgroup of $\pi_n(Y)$.

THEOREM 2. *Let K be a connected finite or infinite cell-complex and Y an arcwise connected topological space such that $\pi_i(K) = \pi_i(Y) = 0$ for $i < n$. If $n = 1$ this last condition is replaced by the condition that K is 1-simple and Y is i-simple for $i = 1, 2, \cdots, \dim K$.*

(2.1) *If ${}^{i+1}H_{\pi_i(Y)}(K) = 0$ for $i = n+1, n+2, \cdots$ then given a homomorphism h mapping $\pi_n(K)[= {}^n\mathcal{H}^I(K)]$ into a subgroup of $\pi_n(Y)[= {}^n\mathcal{H}^I(Y)]$ there is an $f \epsilon Y^K$ such that $h_f^n = h$.*

(2.2) *If ${}^iH_{\pi_i(Y)}(K) = 0$ for $i = n+1, n+2, \cdots$ then given $f_0, f_1 \epsilon Y^K$ we have $h_{f_0}^n = h_{f_1}^n$ if and only if $f_0 \simeq f_1$.*

(2.3) *If ${}^iH_{\pi_i(Y)}(K) = {}^{i+1}H_{\pi_i(Y)}(K) = 0$ for $i = n+1, n+2, \cdots$ then the homotopy classes Φ of Y^K are in a (1-1)-correspondence with the homomorphic mappings of $\pi_n(K)[= {}^n\mathcal{H}^I(K)]$ into subgroups of $\pi_n(Y)[= {}^n\mathcal{H}^I(Y)]$. The correspondence is determined by the operation h_Φ^n.*

PROOF. Let $k_0 \epsilon K^{n-1}$. Consider the mapping $1 \epsilon K^K$ such that $1(x) = x$ for every $x \epsilon K$. According to (1.1) there is a $g \epsilon K^K$ such that $g \simeq 1$ and $g(K^{n-1}) = k_0$. According to (1.2) this leads to a cocycle

$$d^n(g) = \sum_i d(g, \sigma_i^n)\sigma_i^n$$

in K with coefficients in $\pi_n(K)$.

Suppose now that conditions of (2.1) are satisfied and let h be a homomorphic mapping of $\pi_n(K)$ into a subgroup of $\pi_n(Y)$. Consider the cocycle

$$d^n = \sum_i h[d(g, \sigma_i^n)]\sigma_i^n$$

in K with coefficients in $\pi_n(Y)$. By (1.5) of Th. 1 there is an $f \epsilon Y^K$ such that $f(K^{n-1}) = y_0$ and $d^n(f) = d^n$. We are going to prove that $h_f^n = h$.

Since $\pi_n(K)$ and ${}^n\mathcal{H}^I(K)$ are identical, every element $a \epsilon \pi_n(K)$ can be represented as a finite cycle

$$\sum_i \alpha_i\sigma_i^n$$

where α_i are integers and only a finite number of them are $\neq 0$. Using the definition of $d(g, \sigma_i^n)$ and of $d(f, \sigma_i^n)$ it follows that

$$h_g^n(a) = \sum_i \alpha_i d(g, \sigma_i^n), \qquad h_f^n(a) = \sum_i \alpha_i d(f, \sigma_i^n)$$

[4] *Proc. Akad. Amsterdam* 38 (1935), p. 521.

But since $d(f, \sigma_i^n) = h[d(g, \sigma_i^n)]$ it follows that

$$h_f^n(a) = h[h_g^n(a)].$$

Since $g \simeq 1$ we have $h_g^n(a) = a$; therefore $h_f^n(a) = h(a)$ and $h_f^n = h$. This proves (2.1).

Let $f_0, f_1 \in Y^K$ and let $h_{f_0}^n = h_{f_1}^n$. Without restricting the generality we may assume that $f_0(k_0) = f_1(k_0) = y_0$. Consider the mapping $f_j g \in Y^K$ for $j = 0, 1$. Since $f_j g(K^{n-1}) = y_0$ we may consider the cocycles $d^n(f_j g)$ in K with coefficients in $\pi_n(Y)$. Given any n-cell σ_i^n of K we get, according to the meaning of h_j^n,

$$d(f_j g, \sigma_i^n) = h_j^n[d(g, \sigma_i^n)].$$

This implies $d(f_0 g, \sigma_i^n) = d(f_1 g, \sigma_i^n)$, since $h_0^n = h_1^n$. Therefore $d^n(f_0 g) = d^n(f_1 g)$.

Now, if the hypothesis of (2.2) is satisfied, then by (1.6) of Th. 1 we have $f_0 g \simeq f_1 g$. Since $f_0 g \simeq f_0$ and $f_1 g \simeq f_1$ this gives $f_0 \simeq f_1$. Hence (2.2) is proved.

(2.3) follows immediately from (2.1) and (2.2).

3. Let G be an abelian group. Given a finite cycle

$$a^n = \sum_i \alpha_i \sigma_i^n$$

in K with integer coefficients and an arbitrary cocycle

$$d^n = \sum_i \beta_i \sigma_i^n$$

in K with coefficients in G, we define an element $a^n d^n$ of G by taking

$$a^n d^n = \sum_i \alpha_i \beta_i$$

It is easy to verify that if a^n is the boundary of a finite $(n+1)$-chain or if d^n is the coboundary of a $(n-1)$-chain then $a^n d^n = 0$. Consequently given two elements $\boldsymbol{a}^n \in {}^n\mathcal{H}^I(K)$ and $\boldsymbol{d}^n \in {}^n\boldsymbol{H}_G(K)$ the multiplication $\boldsymbol{a}^n\boldsymbol{d}^n$ is a uniquely defined element of G. For a fixed $\boldsymbol{a}^n$ it gives a homomorphic mapping of ${}^n\boldsymbol{H}_G(K)$ into a subgroup of G. Similarly for fixed $\boldsymbol{d}^n$ it gives a homomorphic mapping of ${}^n\mathcal{H}^I(K)$ into a subgroup of G.

Now, let Y be an arcwise connected topological space such that $\pi_i(Y) = 0$ for $i < n$. If $n = 1$ we assume that Y is 1-simple instead. We choose G to be the group $\pi_n(Y)$ which can be considered identical with ${}^n\mathcal{H}^I(Y)$.

Let $f \in Y^K$ and let $f(K^{n-1}) = y_0$. We shall consider the cocycle

$$d^n(f) = \sum_i d(f, \sigma_i^n)\sigma_i^n$$

and its relation to the homomorphic mapping h_f^n of ${}^n\mathcal{H}^I(K)$ into a subgroup of ${}^n\mathcal{H}^I(Y)$ induced by f.

Given a finite cycle

$$a^n = \sum_i \alpha_i \sigma_i^n$$

in K with integer coefficients, we clearly have

$$h_f^n(a^n) = \sum_i \alpha_i\, d(f, \sigma_i^n)$$

and therefore

$$a^n d^n(f) = h_f^n(a^n).$$

Using (1.1) we find

(3.1) *Given any mapping $f \in Y^K$ we have $\boldsymbol{a}^n\boldsymbol{d}^n(f) = h_f^n(\boldsymbol{a}^n)$ for all $\boldsymbol{a}^n \in {}^n\mathcal{H}^I(K)$.*

Our hypothesis concerning Y is satisfied if $Y = S^n$ is the n-dimensional spherical manifold. The group $\pi_n(Y) = {}^n\mathcal{H}^I(Y)$ is then isomorphic to I. Using Th. 2 and (3.1) we prove

THEOREM 3. *Let K be a finite or infinite connected cell-complex such that $\pi_i(K) = 0$ for $i < n$. If $n = 1$ we replace this condition by the condition that $\pi_1(K)$ is abelian.*

(3.2) *If ${}^{i+1}\boldsymbol{H}_{\pi_i(S^n)}(K) = 0$ for $i = n+1, n+2, \cdots$ then for every homomorphic mapping h of the group ${}^n\mathcal{H}^I(K)$ into a subgroup of I there is a cohomology class $d^n \in {}^n\boldsymbol{H}_I(K)$ such that $\boldsymbol{a}^n\boldsymbol{d}^n = h(\boldsymbol{a}^n)$ for every $\boldsymbol{a}^n {}^n\mathcal{H}^I(K)$.*

(3.3) *If ${}^i\boldsymbol{H}_{\pi_i(S^n)}(K) = {}^{i+1}\boldsymbol{H}_{\pi_i(S^n)}(K) = 0$ for $i = n+1, n+2, \cdots$ then given $\boldsymbol{d}^n \in {}^n\boldsymbol{H}_I(K)$ such that $\boldsymbol{a}^n\boldsymbol{d}^n = 0$ for all $\boldsymbol{a}^n \in {}^n\mathcal{H}^I(K)$ we have $\boldsymbol{d}^n = 0$.*

It follows from (3.2) and (3.3) that if the hypothesis of (3.3) is satisfied then the homomorphisms h mapping the group ${}^n\mathcal{H}^I(K)$ into subgroups of I and the elements $\boldsymbol{d}^n$ of ${}^n\boldsymbol{H}_I(K)$ are in a (1-1)-correspondence. h and $\boldsymbol{d}^n$ correspond to each other if and only if

$$\boldsymbol{a}^n\boldsymbol{d}^n = h(\boldsymbol{a}^n) \text{ for all } \boldsymbol{a}^n \in {}^n\mathcal{H}^I(K).$$

PROOF OF TH. 3. If the hypothesis of (3.2) is satisfied then given the homomorphism h there is by (2.1) of Th. 2 a mapping $f \in S^{n^K}$ such that $h_f^n = h$. By (3.1) we then have $\boldsymbol{a}^n\boldsymbol{d}^n(f) = h(\boldsymbol{a}^n)$ for every $\boldsymbol{a}^n \in {}^n\mathcal{H}^I(K)$. This proves (3.2).

If the hypothesis of (3.3) is satisfied then by (2.1) of Th. 2 there is for every $\boldsymbol{d}^n$ a mapping $f \in Y^K$ such that $\boldsymbol{d}^n(f) = \boldsymbol{d}^n$. If now $\boldsymbol{a}^n\boldsymbol{d}^n = 0$ for all $\boldsymbol{a}^n \in {}^n\mathcal{H}^I(K)$ then it follows from (3.1) that $h_f^n = 0$. By (2.2) of Th. 2 we then have $f \simeq f_0$ where $f_0(K) = y_0$. This implies, by (1.6) of Th. 1, that $\boldsymbol{d}^n(f_0) = \boldsymbol{d}^n(f)$. Therefore $\boldsymbol{d}^n = 0$ and (3.3) is proved.

The following example will show that in Th. 3 the hypothesis that $\pi_i(K) = 0$ for $i < n$ cannot generally be removed if $n > 1$. Let Σ be one of D. van Dantzig's solenoids[5] imbedded in S^3. Let K^3 be a subdivision of the open and connected set $S^3 - \Sigma$ into an infinite cell-complex. It is clear that ${}^i\boldsymbol{H}_G(K^3) = 0$ for $i = 3, 4, \cdots$ and any abelian group G. It has been recently proved by N. E. Steenrod[6] that the 1st homology group of K^3, constructed using infinite cycles and integer coefficients, is not enumerable. Since K^3 is a manifold this implies immediately that ${}^2\boldsymbol{H}_I(K^3)$ is not enumerable. On the other hand since Σ is connected it follows from duality theorems that ${}^2\mathcal{H}^I(K^3) = 0$. Hence (3.3)

[5] *Fund. Math.* 15 (1930), pp. 102–125.

[6] These *Annals*, vol. 41 (1940), pp. 833–851.

does not hold for $K = K^3$ and $n = 2$. Naturally, the condition $\pi_1(K^3) = 0$ is not satisfied.

4. Two topological spaces X and Y are said to have the same *homotopy type*[7] if there are two mappings $f \epsilon Y^X$ and $g \epsilon X^Y$ such that $gf \epsilon X^X$ is homotopic to the mapping $1 \epsilon X^X$ defined by the condition $1(x) = x$ for every $x \epsilon X$, and similarly $fg \epsilon Y^Y$ is homotopic to $1 \epsilon Y^Y$.

THEOREM 4. *Two finite or infinite connected cell-complexes K_1 and K_2 which satisfy the following set of conditions have the same homotopy type:*

(4.1) $\pi_i(K_1) = \pi_i(K_2) = 0$ *for* $i < n$,

(4.2) $\pi_n(K_1)$ *and* $\pi_n(K_2)$ *are isomorphic,*

(4.3) ${}^{i+1}H_{\pi_i(K_2)}(K_1) = {}^{i+1}H_{\pi_i(K_1)}(K_2) = 0$ *for* $i = n + 1, n + 2, \cdots$,

(4.4) ${}^{i}H_{\pi_i(K_1)}(K_1) = {}^{i}H_{\pi_i(K_2)}(K_2) = 0$ *for* $i = n + 1, n + 2, \cdots$.

If $n = 1$ we replace (4.1) by the following

$(4.1)_1$ *K_1 and K_2 are i-simple for $i = 1, 2, \cdots, k$, where $k = \max(\dim K_1, \dim K_2)$.*

PROOF. Let h be an isomorphic mapping of $\pi_n(K_1)$ on $\pi_n(K_2)$ and let h^* be the inverse isomorphism. Applying (2.1) of Th. 2 to $K_2^{K_1}$ we obtain a mapping $f \epsilon K_2^{K_1}$ such that $h_f^n = h$. Similarly we obtain a mapping $g \epsilon K_1^{K_2}$ such that $h_g^n = h^*$. Consequently $h_{gf}^n = h^*h$. It follows that $h_{gf}^n = h_1^n$, where $1 \epsilon K_1^{K_1}$ is defined by $1(x) = x$ for all $x \epsilon K_1$. Applying (2.2) of Th. 2 we find that $gf \simeq 1$. Similarly $fg \simeq 1$ where $1 \epsilon K_2^{K_2}$. This proves that K_1 and K_2 have the same homotopy type.

5. In the following we are going to characterize complexes which have the same homotopy type as S^n. The cases $n > 1$ and $n = 1$ will be treated separately.

THEOREM 5. *A finite or infinite connected cell-complex K has the homotopy type of S^n $(n > 1)$ if and only if*

(5.1) $$\pi_1(K) = 0,$$

(5.2) $${}^{i}\mathcal{H}^I(K) = 0 \textit{ for } 1 < i < n,$$

(5.3) $${}^{n}\mathcal{H}^I(K) \textit{ is cyclic infinite},$$

(5.4) $${}^{i+1}H_{\pi_i(S^n)}(K) = 0 \text{ for } i = n + 1, n + 2, \cdots,$$

(5.5) $${}^{i}H_{\pi_i(K)}(K) = 0 \text{ for } i = n + 1, n + 2, \cdots.$$

THEOREM 5_1. *A finite or infinite connected cell-complex K has the homotopy type of S^1 if and only if*

(5.6) $$\pi_1(K) \textit{ is cyclic infinite},$$

(5.7) $$K \textit{ is i-simple for } i = 2, 3, \cdots, \dim K,$$

(5.8) $${}^{i}H_{\pi_i(K)}(K) = 0 \textit{ for } i = 2, 3, \cdots.$$

[7] W. Hurewicz, *Proc. Akad. Amsterdam* 39 (1936), p. 124.

Proofs. Sufficiency. $n > 1$. By a theorem of Hurewicz[4] it follows from (5.1) and (5.2) that $\pi_i(K) = 0$ for $i < n$. Therefore

$$\pi_i(K) = \pi_i(S^n) = 0 \text{ for } i < n. \tag{5.9}$$

From the same theorem it follows that $\pi_n(K)$ and ${}^n\mathcal{H}^I(K)$ are isomorphic. Therefore

$$\pi_n(K) \text{ and } \pi_n(S^n) \text{ are isomorphic.} \tag{5.10}$$

Since S^n is n-dimensional therefore

$${}^iH_G(S^n) = 0 \text{ for } i = n + 1, n + 2, \cdots \text{ and any } G. \tag{5.11}$$

From (5.4), (5.5), (5.9), (5.10) and (5.11) it follows that the conditions of Th. 4 are satisfied. Therefore K and S^n have the same homotopy type.

$n = 1$. Since $\pi_1(K)$ is cyclic infinite therefore we see that

$$\pi_1(K) \text{ and } \pi_1(S^1) \text{ are isomorphic.} \tag{5.12}$$

Since S^1 is i-simple for every i and $\pi_1(K)$ is abelian it follows from (5.7) that

$$K \text{ and } S^1 \text{ are } i\text{-simple for } i = 1, 2, \cdots, \dim K. \tag{5.13}$$

Since S^1 is 1-dimensional and $\pi_i(S^1) = 0$ for $i > 1$ it follows that

$${}^iH_G(S^1) = 0 \text{ for } i = 2, 3, \cdots \text{ and any } G, \tag{5.14}$$

$${}^{i+1}H_{\pi_i(S^1)}(K) = 0 \text{ for } i = 2, 3, \cdots. \tag{5.15}$$

From (5.8) and (5.12)–(5.15) it follows that the conditions of Th. 4 are satisfied and therefore K and S^1 have the same homotopy type.

The necessity follows from the following two lemmas:

(5.16) *If X and Y are arcwise connected and have the same homotopy type then $\pi_i(X)$ is isomorphic with $\pi_i(Y)$, and X is i-simple if and only if Y is.*

(5.17) *If the finite or infinite connected cell-complexes K_1 and K_2 have the same homotopy type then ${}^iH_G(K_1)$ is isomorphic to ${}^iH_G(K_2)$, and ${}^i\mathcal{H}^G(K_1)$ is isomorphic to ${}^i\mathcal{H}^G(K_2)$.*

The proofs are left out.

Since $\pi_i(K) = 0$ implies that K is i-simple it follows that if $\pi_1(K)$ is cyclic infinite and $\pi_i(K) = 0$ for $i > 1$, then the conditions of Th. 5_1 are satisfied and K has the homotopy type of S^1. This is a theorem of Hurewicz.[8]

6. Theorem 5 leads to the following characterization of the finite complexes which have the homotopy type of S^n.

Theorem 6. *A finite polyhedron K has the homotopy type of S^n if and only if K has the same fundamental group and the same homology groups (integer*

---

[8] *Proc. Akad. Amsterdam* 39 (1936), p. 221.

[9] If k is n-dimensional, this is a theorem of Hurewicz, loc. cit. In the general case it also follows from some unpublished results of Hurewicz.

coefficients) as S^n.[9] *If* $n = 1$ *there is an additional condition that* K *is to be i-simple for* $i = 2, 3, \cdots, \dim K$.

PROOF. The necessity follows from (5.16) and (5.17). If K has the same integer coefficient homology groups as S^n then since K is finite it follows that ${}^iH_G(K)$ is isomorphic with ${}^iH_G(S)$ for any i and any coefficient group G. This shows that if the conditions of Th. 6 are satisfied then the conditions of Th. 5 (or Th. 5_1) are satisfied too, and consequently K has the homotopy type of S^n.

There is no example known to prove that the additional condition in the case $n = 1$ is really necessary in Th. 6.

7. Theorem 4 will be used again to prove the following theorem concerning homotopy groups.

THEOREM 7. *Let* K^{n+1} *be an at most* $(n+1)$*-dimensional* $(n > 1)$ *finite or infinite connected cell-complex such that* $\pi_i(K^{n+1}) = 0$ *for* $i < n$. *If the group* $\pi_n(K^{n+1})$ *is a free group with a finite number* $k > 0$ *of generators then* $\pi_{n+1}(K^{n+1}) \neq 0$.

PROOF. Suppose that $\pi_{n+1}(K^{n+1}) = 0$, then

$$ {}^{n+1}H_{\pi_{n+1}(K^{n+1})}(K^{n+1}) = 0. \tag{7.1} $$

Since K^{n+1} is at most $(n+1)$-dimensional we have

$$ {}^iH_G(K^{n+1}) = 0 \text{ for } i > n+1 \text{ and any } G. \tag{7.2} $$

Let K_1^n be a complex obtained by taking k copies of S^n disjoint except for one point which is common for all of them. It is clear that

$$ \pi_i(K^{n+1}) = \pi_i(K_1^n) = 0 \text{ for } i < n, \tag{7.3} $$

$$ \pi_n(K^{n+1}) \text{ and } \pi_n(K_1^n) \text{ are isomorphic}, \tag{7.4} $$

$$ {}^iH_G(K_1^n) = 0 \text{ for } i > n \text{ and any } G. \tag{7.5} $$

It follows from (7.1)–(7.5) that Th. 4 can be applied to demonstrate that K^{n+1} and K_1^n have the same homotopy type. Since $\pi_{n+1}(S^n) \neq 0$ for $n > 1$[10] it follows that $\pi_{n+1}(K_1^n) \neq 0$ and from (5.16) that $\pi_{n+1}(K^{n+1}) \neq 0$.

Th. 7 can also be proved in the case when $\pi_n(K^{n+1})$ is the unrestricted direct product of $\aleph_0$ cyclic infinite groups.

8. In the following we shall discuss the situation arising when S^{r-n-1} is topologically imbedded into S^r.

THEOREM 8. *Let* S_1^{r-n-1} *be a homeomorphic image of* S^{r-n-1} *contained in* S^r *where* $r > n > 0$.

If $n > 1$ *then* $S^r - S_1^{r-n-1}$ *has the homotopy type of* S^n *if and only if* $\pi_1(S^r - S_1^{r-n-1}) = 0$.

---

[10] L. Pontrjagin, *C. R. Acad. Sci. URSS* 19 (1938), p. 147; H. Freudenthal, *Comp. Math.* 5 (1937), p. 301.

If $n = 1$ then $S^r - S_1^{r-2}$ has the homotopy type of S^1 if and only if the group $\pi_1(S^r - S_1^{r-2})$ is cyclic infinite and $S^r - S_1^{r-2}$ is i-simple for $i = 2, 3, \cdots r$.

PROOF. The necessity of the conditions follows from (5.16). In order to prove that they are also sufficient note that it follows from duality theorems that

(8.1) $^i\mathcal{H}^G(S^r - S_1^{r-n-1})$ and $^i\mathcal{H}^G(S^n)$ are isomorphic for $i = 0, 1, \cdots$ and any G.

Since $S^r - S_1^{r-n-1}$ is a manifold the group $^iH_G(S^r - S_1^{r-n-1})$ is isomorphic with the $(r - i)$-th homology group of $S^r - S_1^{r-n-1}$ obtained using infinite chains. Since S_1^{r-n-1} is locally connected in all dimensions it follows from a recent theorem of Steenrod[6] that this group is isomorphic with $^{r-i-1}\mathcal{H}^G(S_1^{r-n-1})$. Therefore

(8.2) $^iH_G(S^r - S_1^{r-n-1}) = 0$ for $i = n + 1, n + 2, \cdots$ and any G.

Now, (8.1) and (8.2) and the conditions of Th. 8 permit us to apply Th. 5 (or Th. 5_1) and prove that $S^r - S_1^{r-n-1}$ has the homotopy type of S^n.

9. A set $Y \subset X$ is called a *retract* of X if there is a mapping $r \in X^X$ (called a *retraction*) such that $r(X) = Y$ and $r(y) = y$ for all $y \in Y$. If $r \simeq 1$, where $1 \in X^X$ is defined by $1(x) = x$ for all $x \in X$, then Y is called a *deformation retract* of X.

We shall consider in S^r a homeomorphic image S_0^n of S^n and a homeomorphic image S_1^{r-n-1} of S^{r-n-1}, where $r > n \geqq 0$ and $S_0^n S_1^{r-n-1} = 0$. Assigning orientations to S_0^n and S_1^{r-n-1} we obtain then a linkage coefficient whose absolute value, denoted by $\bar{v}(S_0^n, S_1^{r-n-1})$, is independent of the chosen orientations. We have proved in another paper[11] that S_0^n is a retract of $S^r - S_1^{r-n-1}$ if and only if $\bar{v}(S_0^n, S_1^{r-n-1}) = 1$. We shall now discuss the analogous problem obtained by replacing retracts by deformation retracts.

THEOREM 9. *Let S_0^n and S_1^{r-n-1} be disjoint homeomorphic images of S^n and S^{r-n-1} contained in S^r, where $r > n > 0$.*

If $n > 1$ then S_0^n is a deformation retract of $S^r - S_1^{r-n-1}$ if and only if $\bar{v}(S_0^n, S_1^{r-n-1}) = 1$ and $\pi_1(S^r - S_1^{r-n-1}) = 0$.

If $n = 1$ then S_0^1 is a deformation retract of $S^r - S_1^{r-2}$ if and only if $\bar{v}(S_0^1, S_1^{r-2}) = 1$, $\pi_1(S^r - S_1^{r-2})$ is cyclic infinite, and if $S^r - S_1^{r-2}$ is i-simple for $i = 2, 3, \cdots, r$.

PROOF. If S_0^n is a deformation retract of $S^r - S_1^{r-n-1}$ then by the theorem quoted above $\bar{v}(S_0^n, S_1^{r-n-1}) = 1$. It also follows that S_0^n and $S^r - S_1^{r-n-1}$ have the same homotopy type and therefore the remaining conditions of Th. 9 are fulfilled in view of Th. 8.

Now, let $\bar{v}(S_0^n, S_1^{r-n-1}) = 1$ and let r be a retraction mapping $S^r - S_1^{r-n-1}$ into S_0^n. Let γ_0^n be a basis-cycle for the group $^n\mathcal{H}^I(S_0^n)$. We shall prove that γ_0^n is also a basis-cycle for $^n\mathcal{H}^I(S^r - S_1^{r-n-1})$. In fact, if this is not the case there is a singular cycle γ^n in $S^r - S_1^{r-n-1}$ such that $\gamma_0^n \sim k\gamma^n$ in $S^r - S_1^{r-n-1}$

---

[11] *Fund. Math.* 31 (1938), p. 192, see also K. Borsuk and S. Eilenberg, *Fund. Math.* 26 (1936), p. 215.

where $k \neq \pm 1$. This implies $\gamma_0^n = r(\gamma_0^n) \sim kr(\gamma^n)$ in S_0^n contradicting our hypothesis that γ_0^n was a basis-cycle for ${}^n\mathcal{H}^I(S_0^n)$. Therefore γ_0^n is a basis-cycle for ${}^n\mathcal{H}^I(S^r - S_1^{r-n-1})$. Since $r(\gamma_0^n) = \gamma_0^n$ it follows that $h_r^n = h_1^n$, where 1 is the identity mapping of $S^r - S_1^{r-n-1}$ into itself.

If the remaining conditions of Th. 9 are also fulfilled then, by Th. 8, $S^r - S_0^{r-n-1}$ has the homotopy type of S_0^n. From Th. 5 (or Th. 5_1) it then follows that (2.2) of Th. 2 can be applied to mappings of $S^r - S_1^{r-n-1}$ into itself. Since $h_r^n = h_1^n$, we have $r \simeq 1$ so that S_0^n is a deformation retract of $S^r - S_0^{r-n-1}$.

10. If we take $n = 0$ in Th. 8 and Th. 9 then $S^r - S_1^{r-1}$ is not connected but consists of two connected regions C_1 and C_2. By duality theorems we have ${}^i\mathcal{H}^I(C_j) = 0$ for $i = 1, 2, \cdots$ and $j = 1, 2$. If we further admit that $\pi_1(C_j) = 0$ then by a theorem of Hurewicz[4] C_j can be deformed to a point. It follows that $S^r - S_1^{r-1}$ has then the homotopy type of S^0 where S^0 consists of two points. It is also clear that then every $S_0^0 \subset S^r - S_1^{r-1}$ such that $\bar{v}(S_0^0, S_1^{r-1}) = 1$ is a deformation retract of $S^r - S_1^{r-1}$. Hence we see that

(10.1) *The statements of Th. 8 and Th. 9 concerning the case $n > 1$ hold also for $n = 0$ provided the relation $\pi_1(S^r - S_1^{r-1}) = 0$ is interpreted as $\pi_1(C_1) = \pi_1(C_2) = 0$ where C_1 and C_2 are the components of $S^r - S_1^{r-1}$.*

In the case $n = 1$ Theorems 8 and 9 contain the condition that $S^r - S_1^{r-2}$ should be i-simple for $i = 2, 3, \cdots, r$. In the case $r = 3$ (and also in the trivial case $r = 2$) the author has proved[12] that this condition can be dropped. The similar question concerning $r > 3$ remains unanswered and seems to be closely related to the following problem:

Let S_1^{r-n-1} be a homeomorphic image of S^{r-n-1} in S^r where $r > n > 0$. Under what conditions is $S^r - S_1^{r-n-1}$ aspherical (i.e. $\pi_i(S^r - S_1^{r-n-1}) = 0$ for $i > 1$)? In particular, is $S^r - S_1^{r-n-1}$ aspherical if $n = 1$ and $\pi_1(S^r - S_1^{r-2})$ is cyclic infinite?

Again this question has been answered positively only for $r = 2, 3$.[12]

UNIVERSITY OF MICHIGAN

[12] *Fund. Math.* 28 (1937), p. 238.

Annals of Mathematics
Vol. 42, No. 2, April, 1941

ON PROJECTIVE EQUIVALENCE OF TRILINEAR FORMS

By Robert M. Thrall

(Received March 20, 1940)

1. Introduction. The two fundamental problems in the projective classification of trilinear forms are: (I) the determination of necessary and sufficient conditions for the (projective) equivalence[1] of two trilinear forms; (II) the determination of a set of forms which consists of one and just one representative form from each *class* of equivalent forms.

In the form[2] $F(x, y, z) = a_{\alpha\beta\delta}y_\alpha z_\beta x_\delta$ let there be py's, qz's, and rx's and suppose that the numbers p, q, r are the two way rank invariants[3] of the form F. Let the variables be so named that $p \leqq q \leqq r$. We can measure progress in the solution of problems I and II in terms of the sets of invariants (p, q, r) for which they have been solved.

Problems I and II have been rationally solved (i.e. for the coefficients $a_{\alpha\beta\delta}$ belonging to any field) for $p = 1$ and $p = 2$. For if $p = 1$, then $q = r$ and is the only invariant of F. For $p = 2$ the trilinear form can be considered as a pencil of bilinear forms. Two such pencils are equivalent if and only if they have the same minimal numbers and projectively equivalent invariant factors.[4]

If $p > 2$ problems of rationality enter, and most of the progress here has been for forms with coefficients in the field, K, of complex numbers; although in [13], the case $(p, q, r) = (3, 3, 3)$ was discussed for forms belonging to a Galois field by rational methods applicable to any field. [14] contains the complete solution for forms $(3, 3, 3)$ with coefficients in K.

In §§2–9 of the present paper, methods similar to those used in [14] are applied to the next case $(3, 3, 4)$. The coefficient field is K although many of the results are obviously more general. §2 is devoted to definitions and notation for determinantal manifolds associated with trilinear forms. In §3 the forms $(3, 3, 4)$ are divided into generate and non-degenerate cases according to their associated determinantal manifolds. §§4–8 give classification theorems for the various non-degenerate cases. Theorem 4.3 is basic in this discussion. §5 discusses briefly generalization of 4.3 to forms $(p, p, 2p - 2)$. The principal results of these sections are found in theorems 6.5, 6.6, 7.1, 8.6. §9 (especially theorem 9.1) and the appendix contain the solutions of problems I and II for the degen-

---

[1] For definitions of equivalence and ρ-equivalence, see [14] p. 678. (The numbers in brackets refer to the bibliography.)

[2] Repeated Greek indices indicate summation.

[3] This means that F is equivalent to no form with less than py's, qz's and rx's. For discussion of and reference to these invariants, see [14] p. 678.

[4] See [9], and [12] p. 682.

erate cases. In §10 the work of T. G. Room[5] on determinantal manifolds is applied to the "general" case for $3 \leqq p < q \leqq r$ and $3 < p = q \leqq r$ to obtain theorems which reduce problem I to a study of projective equivalence of manifolds related to the forms. Also in the section is a discussion of different interpretations of the concept "general" as applied to trilinear forms.

2. The determinantal manifolds[6] associated with a trilinear form. Let $F(x, y, z) = a_{\alpha\beta\delta}y_\alpha z_\beta x_\delta \equiv y_\alpha z_\beta x_{\alpha\beta} \equiv z_\beta x_\delta y_{\beta\delta} \equiv y_\alpha x_\delta z_{\alpha\delta}$ define the elements of the matrices $\| x_{\alpha\beta} \|$, $\| y_{\beta\delta} \|$, $\| z_{\alpha\delta} \|$. Denote by $V_x^s = (| p, q |_s[r - 1])$ the manifold consisting of points x in projective $r - 1$ space for which $\| x_{\alpha\beta} \|$ is of rank s or less. We write the equations for V_x^s in the form $| x_{\alpha\beta} |_s = 0$. We call $V_x = V_x^{p-1}$ *the principal manifold* of the matrix $\| x_{\alpha\beta} \|$ and of the form F.[7] The manifolds V_y^s and V_z^s are defined analogously from the matrices $\| y_{\beta\delta} \|$ and $\| z_{\alpha\delta} \|$. The equivalence of two trilinear forms obviously implies the projective equivalence of their corresponding manifolds[8] V, and in general the converse is also true. (See §10.)

Since the derivative of an n-rowed determinant is a linear combination of its $n - 1$ rowed minors, the manifold V^{s-1} must be multiple on V^s. In general V_x^s has dimension $r - 1 - (q - s)(p - s)$ and so will necessarily be non vacuous only for values of s for which this dimension is positive. In particular cases, however, the dimension may be greater than this number.

3. Trilinear forms (3, 3, 4), **general theory.** In general, $V_x = V_x^2$ (the principal manifold), a cubic surface in 3-space, and V_y^2 and V_z^2, zero dimensional manifolds of order six in 2-space, are the only non empty manifolds associated with a trilinear form (3, 3, 4). However, there are degenerate cases in which some of the V^1's exist, or in which V_y^2 or V_z^2 is of dimension greater than zero. On this basis we divide our treatment into two parts:

(A) *The non-degenerate cases*: V_y^1 is empty, and V_y^2 is of dimension zero and order six.

(B) *The degenerate cases*: At least one of the conditions under (A) is not satisfied.

Before treating (A) and (B) separately we study three algebraic transformations defined by the form F. We say that $\bar{y}$ and $\bar{z}$ are images under the transformation T_1 if

$$a_{\alpha\beta\delta}\bar{y}_\alpha\bar{z}_\beta = 0, \qquad \delta = 1, 2, 3, 4. \tag{3.1}$$

The transformations T_2 and T_3 are similarly defined by

$$a_{\alpha\beta\delta}\bar{z}_\beta\bar{x}_\delta = 0, \qquad \alpha = 1, 2, 3, \qquad \text{and} \qquad a_{\alpha\beta\delta}\bar{y}_\alpha\bar{x}_\delta = 0, \qquad \beta = 1, 2, 3.$$

[5] [3] especially Chapter VII, or [10]. Chapter VII is based upon [10]. For a review of [3] see [11].

[6] The definitions and notation are those given by Room [3] Chapter I.

[7] If $r = q$ $(r = q = p)$ F has two (three) principal manifolds, but if $r > q \geqq p$ there is just one principal manifold.

[8] See [13] p. 386.

Considered as linear equations in $\bar{z}$ the number of solutions of (3.1) depends upon the rank of $\| y_{\beta\delta} \|$, and inversely for $\bar{y}$ and $\| z_{\alpha\delta} \|$. Thus if $\bar{y}$ and $\bar{z}$ are images under T_1, then $\bar{y} \in V_y^2$ and $\bar{z} \in V_z^2$. If $\bar{y} \in V_y^1$ it will have a whole line of images $\bar{z}$, which must all lie on V_z^2. If $\bar{y} \in V_y^2$ but $\notin V_y^1$, then it will have a single image $\bar{z}$. Hence, if V_y^2 is of dimension zero, V_z^1 is empty; and, if further V_y^1 is empty, the points of V_y^2 and V_z^2 will be in 1-1 correspondence.

From this it follows that the conditions under (A) above are actually symmetric in y and z although stated in terms of y only. A zero dimensional manifold, V_y^2, of order six is, in general, a set of six distinct points. However, these points may coincide in various ways so that the actual number of distinct points in V_y^2 may be anything from one to six. The above arguments show that *under* (A) *if V_y^2 has k distinct points then V_z^2 likewise has k distinct points.* We shall show later that $\bar{y}$ has the same multiplicity in V_y^2 as its image $\bar{z}$ has in V_z^2.

T_2 is a transformation between points of V_z^3 (the whole z-plane) and points of V_x^2. If $\bar{z} \in V_z^1$ its image under T_2 is a plane, which must be a linear factor of V_x^2; if $\bar{z} \in V_z^2$ but $\notin V_z^1$ its image is a line on V_x^2; and if $\bar{z} \notin V_z^2$, its image is a single point of V_x^2. T_3 is of the same general character as T_2.

4. The non-degenerate cases. Consider the general trilinear form (3, 3, 4) in which the coefficients $a_{\alpha\beta\delta}$ are independent indeterminants. The three rowed minors of $\| y_{\beta\delta} \|$ define a web, W_y, of cubics whose base is the six (in this case distinct) points of V_y^2. Let $f_\delta(y)$ denote the three rowed minor of $\| y_{\beta\delta} \|$ obtained by dropping the δ-th column. Then two generic members $G(\lambda, y) = \lambda_\delta f_\delta(y)$ and $G(\mu, y)$ of W_y will meet in nine points of which six, say $y^{(1)}, \cdots, y^{(6)}$ have coordinates free of λ and μ. The u-resultant[9] of $G(\lambda, y)$ and $G(\mu, y)$ will therefore split rationally into a factor

$$H(u, a_{\alpha\beta\delta}) = \prod_{\gamma=1}^{6} y_\alpha^{(\gamma)} u_\alpha$$

of degree 6 and free of λ and μ, and a second factor of degree 3.

If now we give the $a_{\alpha\beta\delta}$ arbitrary values in the ground field K, for which $H(u, a_{\alpha\beta\delta}) \not\equiv 0$, some or all of the six points, $y^{(\gamma)}$, may coincide. If a point, P, of V_y^2 arises from an s-fold factor of H then we assign to P the multiplicity s in V_y^2, thus ensuring that V_y^2 will have multiplicity 6 in any form for which $H \not\equiv 0$. On the other hand this multiplicity s assigned to P is by its definition the same as the multiplicity of intersection at P of two generic cubics of the web W_y.

We now divide the non-degenerate cases into two sets:

The linear non-generate cases: W_y contains a curve of genus one.

The nodal non-degenerate cases: W_y contains no curve of genus one. In this section we shall treat only the linear cases.

If C is a curve of genus one belonging to W_y, we can utilize its Puiseux ex-

---

[9] See [5] vol. 2 p. 21.

pansions to characterize V_y^2. To obtain uniqueness we choose the particular expansion

$$y_\alpha = E_\alpha(t) \quad \alpha = 1, 2, 3$$

at a point $P = (c_1, c_2, c_3)$ for which

$$E_i(t) = 1, \quad E_j(t) = c_j/c_i + t, \quad E_k(t) = c_{k\gamma}t^\gamma \quad \gamma = 0, \cdots, \infty \qquad (c_{k0} = c_k/c_i)$$

where i is the smallest index for which $c_i \neq 0$, and $j < k$ unless a vertex of the coordinate triangle lies on the tangent to C at P, in which case $c_{k1} = 0$ and it may be necessary to take $j > k$. That this is possible follows from the linearity of every branch of C and the general theory of Puiseux expansions.[10] We shall call these specialized expansions the *canonical expansion* of C at P.

Let the web, W_y, be given by

$$G(\lambda, y) = \lambda_\delta f_\delta(y) \quad \delta = 1, \cdots, 4$$

If $G(\lambda, E(t)) \equiv 0 \bmod t^s$ but $\not\equiv 0 \bmod t^{s+1}$ then P counts s times in V_y^2. Let $P_\alpha(t)$ be obtained from $E_\alpha(t)$ by dropping all terms whose degree in t is s or greater. If C' is any other member of W_y (of genus one), then the first s terms of its canonical expansion $E'_\alpha(t)$ will likewise be $P_\alpha(t)$ and so the polynomials $P_\alpha(t)$ express properties of the web and not merely of the particular curve C in the web. We call the $P_\alpha(t)$ the *canonical polynomials* of W_y at P. (We note that these polynomials are identically zero unless $P \in V_y^2$.)

Let P^μ, $\mu = 1, \cdots, h$, count s_μ times in V_y^2, $(s_1 + \cdots + s_h = 6)$, and let the canonical polynomials of W_y at P^μ be $P^\mu_\alpha(t)$.

4.1 Definition. The curve $K: g(y) = 0$ is said to pass through V_y^2 if and only if $g(P^\mu(t)) \equiv 0 \bmod t^{s_\mu}$ $\mu = 1, \cdots, h$.

4.2 Theorem. *The cubic $g(y) = 0$ belongs to W_y if and only if it passes through V_y^2.*

For there is just one web of cubics satisfying the six conditions of definition 4.1. (4.2 does not hold in the nodal cases.)

The results of this section obviously hold for the z-space as well as the y-space. The following theorem is fundamental in the solution of problem II for the linear case and with a slight modification applies also to linear branches of V_y^2 in the non linear case.

4.3 Theorem. *Let F be a linear non-degenerate form and let T_1 send P on V_y^2 into P' on V_z^2. If a branch of V_y^2 at P has multiplicity σ, with canonical polynomials $P_\nu(t)$, then there exists a linear branch of V at P' of multiplicity σ with canonical polynomials $P'_\nu(\tau)$, and a polynomial $\tau = \tau(t) = b_1t + b_2t^2 + \cdots + b_{\sigma-1}t^{\sigma-1}$, $b_1 \neq 0$, such that*

$$(B_\sigma): \; F(t, x_{\alpha\beta}) = P_\alpha(t)P'_\beta(\tau(t))x_{\alpha\beta} \equiv 0 \bmod t^\sigma.$$

[10] See for instance [2] pp. 213–231.

PROOF. We take $P = (c_1, c_2, c_3)$, $P' = (c_1', c_2', c_3')$ and for convenience suppose $c_1c_1' \neq 0$.

The four congruences, B_σ, can be regarded as homongeneous linear equations with coefficients in the ring, R, of polynomials in t reduced modulo t^σ. The matrix $\| y_{\beta\delta}P_\alpha(t) \|$ is by hypothesis of rank 2 (in R). Hence there exists a solution $z_\beta = Q_\beta'(t)$ with the following properties:[11] $z_1 = 1$, $z_2 = c_2'/c_1' + E_2(t)$, $z_3 = c_3'/c_1' + E_3(t)$ where the $E_\beta(t)$ belong to R (i.e. are polynomials in t of degree less than σ).

The polynomials $P_\alpha(t)$, not all zero divisors ($P_1(t) = 1$), are solutions of the congruences $z_{\alpha\delta}(Q_\beta'(t))y_\alpha = 0$. Hence the rank of $\| z_{\alpha\delta}(Q_\beta'(t) \|$ is <3 (actually 2 since the case is non-degenerate). To complete the proof of the theorem we need merely to show the existence of a $\tau = b_1t \cdots + b_{\sigma-1}t^{\sigma-1}$, $b_1 \neq 0$ such that $Q_\beta'(t) = P_\beta'(\tau(t))$ where the $P_\beta'(\tau)$ are "canonical" polynomials for V_z^2. If $E_2(t) \not\equiv 0 \bmod t$ we take $\tau = E_2(t)$. Otherwise take $\tau = E_3(t)$. [If both $E_2(t)$ and $E_3(t)$ were divisible by t^2 the whole line $y_k - c_{k1}y_j = (c_k - c_jc_{k1})y_1$ would map into P' under T_1 contrary to the hypothesis of non-degeneracy.]

5. Linear non-degenerate forms $(p, p, 2p - 2)$. The forms (3, 3, 4) belong to the series of forms $(p, p, 2p - 2)$ in which the principal manifold is a determinant, and in which V_y^{p-1} and V_z^{p-1} are, in general, zero dimensional, of order $N = \binom{2p-2}{p-1}$. A form $(p, p, 2p - 2)$ is said to be *non-degenerate* if V_y^{p-1} is zero dimensional of order N, and V_y^{p-2} is empty. A non-degenerate form $(p, p, 2p - 2)$ is said to be *linear* if V_y^{p-1} can be defined by canonical polynomials satisfying 4.1 with the word "curve" replaced by "hypersurface." With these definitions theorem 4.3 is valid for forms $(p, p, 2p - 2)$ if we replace V_y^2, V_z^2 by V_y^{p-1}, V_z^{p-1} throughout. The proof of this generalization is stepwise parallel to that of 4.3.

6. The general non-degenerate case. In the general case V_y^2 consists of six distinct points $y^{(1)}, \cdots, y^{(6)}$, not on a conic and no three on a line. This implies the same for V_z^2, since otherwise T_3 and T_2 could not map the y and z planes respectively into the same surface V_x^2. Let coordinates be so chosen that $y^{(1)} = (1, 0, 0)$, $y^{(2)} = (0, 1, 0)$, $y^{(3)} = (0, 0, 1)$, $y^{(4)} = (1, 1, 1)$, $y^{(5)} = (e, a, b)$, $y^{(6)} = (e, c, d)$; $z^{(1)} = (1, 0, 0)$, $z^{(2)} = (0, 1, 0)$, $z^{(3)} = (0, 0, 1)$, $z^{(4)} = (1, 1, 1)$, $z^{(5)} = (e, a', b')$, $z^{(6)} = (e, c', d')$; and so that

$$y_\alpha^{(i)} z_\beta^{(i)} x_{\alpha\beta} \equiv 0 \qquad i = 1, \cdots, 6 \tag{6.1}$$

expresses the congruence of V_y^2 and V_z^2 under T_1. Of these six equations, one, say the last, must be dependent on the other five; for otherwise there would be

[11] The existence theorem for linear equations in a commutative ring assures the existence of a solution with not all Q_β' null divisors. By taking $\sigma = 1$ we see that Q_1' is not a null divisor (since $c_1' \neq 0$). But every non-null divisor in R has an inverse in R, so that $z_\beta = Q_\beta'/Q_1'$ is also a solution and has $z_1 = 1$.

only three independent $x_{\alpha\beta}$, whereas $r = 4$. This means that the coordinates of $y^{(6)}$ and $z^{(6)}$ are rational functions of the coordinates of $y^{(i)}$, $z^{(i)}$, $i < 6$.

$a \neq b$, for if $a = b$ then $y^{(1)}$, $y^{(4)}$, $y^{(5)}$ are collinear. Hence we can solve the first five equations of 6.1 for x_{21} and x_{31} giving

$$x_{11} = x_{22} = x_{33} = 0, \quad x_{21} = \alpha_1 x_{12} + \alpha_2 x_{13} + \alpha_3 x_{23} + \alpha_4 x_{32},$$

$$x_{31} = (-1 - \alpha_1)x_{12} + (-1 - \alpha_2)x_{13} + (-1 - \alpha_3)x_{23} + (-1 - \alpha_4)x_{32}$$

where

$$\alpha_1 = \frac{b - a'}{a - b}, \qquad \alpha_2 = \frac{b - b'}{a - b}, \qquad \alpha_3 = \frac{eb - ab'}{e(a - b)}, \qquad \alpha_4 = \frac{(e - a')b}{e(a - b)}.$$

If now we set $x_{12} = x_1$, $x_{13} = x_2$, $x_{23} = x_3$, $x_{32} = x_4$ and compute V_y^2 we find

$$y^{(6)} = [e(a - a')(b - b')(a' - b'), (a - a')(b' - e)(ab' - a'b), (b - b')(a' - e)(ab' - a'b)], \tag{6.2}$$

and from the symmetry we obtain $z^{(6)}$ by substituting z, a', b', a, b for y, a, b, a', b' in 6.2.

6.3 Theorem. *Given $y^{(i)}$, $i < 6$, and $z^{(i)}$, $i < 5$ there exists a unique point $z^{(5)}$ for which $y^{(6)}$ is any preassigned point (e, c, d) not on the conic or any of the lines determined by $y^{(i)}$, $i < 6$.*

For proof it is sufficient to observe that the equations

$$\begin{aligned} \rho t_1' &= (at_1 - et_2)(bt_1 - et_3)(t_2 - t_3) \\ \rho t_2' &= (at_1 - et_2)(t_3 - t_1)(at_3 - bt_2) \\ \rho t_3' &= (bt_1 - et_3)(t_2 - t_3)(at_3 - bt_2) \end{aligned} \tag{6.4}$$

define an involutorial plane Cremona transformation. If now we set $t' = y^{(6)}$ and $t = z^{(5)}$, 6.4 reduces to 6.2, so that 6.2 has a unique solution for a', b' in terms of e, a, b, c, d. (The excluded conic and lines include all of the principal curves of the transformation.)

Theorem 6.3 enables us to give the complete solution of problems I and II in the general case. For we have at once the following theorems:

6.5 Theorem. *There exists a trilinear form with $(p, q, r) = (3, 3, 4)$ for which V_y^2 is any preassigned, general set of six points in the y plane.*

6.6 Theorem. *Two forms, F and $\bar{F}$, for which V_y^2 is general are equivalent if and only if V_y^2 is projectively equivalent to $\bar{V}_y^2$.*

The projective invariants of the six point plane manifold have been given by Coble in [7].

6.7 Corollary. *Two non-equivalent general forms F and $\bar{F}$ are g-equivalent if and only if V_y^2 is projectively equivalent to $\bar{V}_z^2$.*

Hence, the g-class containing a form is identical with the class containing the form if and only if V_y^2 is projectively equivalent to V_z^2 (not necessarily in the

order of congruence under T_1). In general this is not the case, but if e, a, b, c, d, satisfy certain syzygies it may happen.

7. The non-general, non-degenerate linear cases. There are approximately a hundred different projective types of special zero dimensional manifolds of order 6. For instance, there may be six distinct points; all on a conic, three on a line, three on one line and the other three on a second line, five on two lines meeting at one of the points, the vertices of a triangle plus a point on each side, the vertices of a quadrilateral.[12] If two of the points coincide (become "infinitely near") along some fixed direction, there are 11 types of ways in which the other four points can be placed relative to the first point and the given direction at it, etc. etc. We shall not consider each such case separately, but the following theorem enables one to determine when two such manifolds U and U' can serve as V_y^2 and V_z^2 for a trilinear form, thus giving the information needed to reduce problem II (for the forms of this section) to a matter of direct computation.

7.1 THEOREM. *If U and U' are two zero dimensional (non-nodal) manifolds of order* 6 *there exists a trilinear form* (3, 3, 4) *with $V_y^2 = U$ and $V_z^2 = U'$ if and only if the following conditions are satisfied.*

a) *the points and branches of U and U' can be put into* 1-1 *correspondence in accordance with the requirements of theorem* 4.3.

b) *the resulting relations B_σ considered as linear equations in the $x_{\alpha\beta}$ shall have a matrix of rank* 5.

c) *the V_z^2 thus defined is irreducible, and without a line of nodes.*

PROOF: Necessity, a) is a consequence of theorem 4.3; b), there are just 9 $x_{\alpha\beta}$ of which just 4 $(=r)$ are independent; c), if V_z^2 were reducible it could not be the rational map of the y plane and so one of its factors would have to be the image of a one dimensional part of V_y^2 or of a point of V_y^1, either possibility giving a degenerate case. If V_z^2 has a line of nodes its plane sections are rational. Hence, the corresponding curves in W_y must be rational; and the case is either degenerate or nodal.

Sufficiency. a) and b) are sufficient to define a form (3, 3, 4) with V_y^2 including U and V_z^2 including U', with the inclusion being equality unless the case is degenerate. Reference to §9 shows that V_z^2 contains a line of nodes in the degenerate cases where it is irreducible. Hence c) implies that the case is non-degenerate (as well as non-nodal).

This theorem is valid for the general case but is weaker than theorems 6.3 and 6.5. The following extension of 6.6, which solves problem I for all non-degenerate, non-nodal cases, is an immediate consequence of 7.1.

7.2 THEOREM. *Two non-degenerate, non-nodal forms F, $\bar{F}$ are equivalent if and only if V_y^2 and V_z^2 are projectively equivalent to $\bar{V}_y^2$ and $\bar{V}_z^2$, respectively.*

[12] If four of the points are on a line, the case is degenerate with the whole line belonging to V_y^2.

The following two lemmas are useful in applications of 7.1.

7.3 LEMMA. *If a line λ cuts V_y^2 3 times in points whose images under T_1 lie on a line λ', then either the case is nodal; or V_y^2 includes λ, V_z^2 includes λ', and the V_x^2 is reducible.* (It is not necessary that λ cut V_y^2 in 3 distinct points.)

7.4 LEMMA. *If V_y^2 lies on a conic, then the case is degenerate.*

Both of the proofs follow upon direct application of 4.3 and will be omitted.

7.1 and 7.3 enable us to prove the following theorem:

7.5 THEOREM. *Not every quaternary cubic can be written as a three rowed determinant whose elements are linear forms.*[13]

More specifically and in the language of this paper the theorem is

7.6 THEOREM. *V_x^2 can be any quaternary cubic excepting one projectively equivalent to $f(x) = x_1^3 + x_2^2x_3 + x_3^2x_4$, characterized projectively by its unode of the third kind.*[14]

PROOF. The surface $f(x) = 0$ contains just one line, μ. Suppose that V_y^2 contains the two points P_1 and P_2 both of which map into μ under T_3. Then each point of μ has two images in the y plane and therefore all of μ lies on V_x^1, i.e. μ is a line of nodes of V_x^2. But $f(x)$ has only one node. We have proved that if V_x^2 is $f(x) = 0$ then V_y^2 has just one point, P. P cannot be a node of W_y for then every plane section of V_x^2 would be rational.

Next suppose that P is not a point of inflexion for all members of the web, W_y. Then both P and the tangent at P to the curves of the web must map into μ under T_3, which would again require μ to be a line of nodes.

The only remaining possibility is that V_y^2 be a single point at which the curves of the mapping web have a common inflexion tangent, λ, and three further "infinitely near" intersections. Similarly, V_z^2 is a single point cut three times by a line λ'. But now λ and λ' satisfy the hypotheses of 7.3, and so by 7.1 no determinantal representation of $f(x)$ can exist. To complete the proof we remark that representations for the other projective types can be readily constructed.

We conclude this section with some applications of 7.1 which illustrate the general method of attack in the special, linear, non-degenerate cases. We ask if W_y can be a web of cubics with a common inflexion tangent at one point and with a triple intersection at a second point; and the same for W_z.

Choose y coordinates so that W_y has the inflexion point (1, 0, 0) with tangent $y_3 = 0$, and so that the second base point is (0, 0, 1) with common tangent $y_1 = 0$. Then the canonical polynomials will be $P^{(1)} = (1, t, 0)$ and $P^{(2)} = (at^2, t, 1)$. Then choose z coordinates so that the canonical polynomials are $Q^{(1)} = (1, \tau_1, b\tau_1^2)$ and $Q^{(2)} = (0, \tau_2, 1)$.

If $Q^{(2)}$ and $P^{(1)}$ were images under T_1, then by 7.3 the case would be degenerate. Hence, $P^{(i)}$ corresponds to $Q^{(i)}$, $i = 1, 2$. We have two sets of equations B_σ,

---

[13] In [8] p. 175 Dickson has proved that every "sufficiently general" quaternary cubic can be written as a determinant, but he gives no counter example. Room ([3] p. 65) states, incorrectly, that any quaternary cubic is determinantal.

[14] See [1] p. 61.

each with $\sigma = 3$. If we can choose constants $b_1 \neq 0, b_2; c_1 \neq 0, c_2$ so that $\tau_1 = b_1 t + b_2 t^2$ and $\tau_2 = c_1 t + c_2 t^2$ we will have met condition a) of 7.1. The equations are:

$$B_3^{(1)} \quad \begin{cases} x_{11} = 0 \\ x_{21} + b_1 x_{12} = 0 \\ b_1 x_{22} + b_2 x_{12} + b b_1^2 x_{13} = 0 \end{cases}$$

$$B_3^{(2)} \quad \begin{cases} x_{33} = 0 \\ x_{23} + c_1 x_{32} = 0 \\ a x_{13} + c_1 x_{22} + c_2 x_{32} = 0. \end{cases}$$

Necessary and sufficient conditions for the dependence of these equations are $b_2 = c_2 = 0$, $c_1 b_1^2 b = a b_1$ or $c_1 = a/bb_1$. This gives

$$\| x_{\alpha\beta} \| = \begin{Vmatrix} 0 & x_{12} & x_{13} \\ -b_1 x_{12} & -bb_1 x_{13} & x_{23} \\ x_{31} & -(a/bb_1)x_{23} & 0 \end{Vmatrix} \sim \begin{Vmatrix} 0 & x_1 & x_2 \\ -x_1 & -x_2 & x_3 \\ x_4 & -x_3 & 0 \end{Vmatrix}.$$

Then $V_x^2 : x_1 x_3 (x_2 + x_4) + x_2^2 x_4 = 0$ is irreducible and has only three nodes so that condition c) of 7.1 is satisfied, and the existence of the form in question is proved. Incidentally, the canonical form obtained shows that there is but one class of forms having V_y^2 and V_z^2 of the above required type.

We can show, further, that for this particular choice of V_y^2 there can be only the one choice of V_z^2 (to within projectivities). For by 7.1 V_z^2 must have just two points, each of multiplicity three. By a purely geometric study of the surface V_x^2 defined by V_y^2, it can be shown that there are just two possible kinds of mapping webs having two base points each of multiplicity 3. One of these is V_y^2 and the other has a base lying on a conic and so is excluded by 7.4.

However, it is not always necessary that V_y^2 and V_z^2 be of the same projective type. For consider the form represented by

$$\| x_{\sigma\beta} \| = \begin{Vmatrix} 0 & x_{12} & x_{13} \\ x_{12} & x_{13} + a x_{12} & x_{31} \\ x_{31} & x_{32} & 0 \end{Vmatrix}.$$

V_y^2 has canonical polynomials $P^{(1)} = (1, t, -t^3)$, $P^{(2)} = (0, t, 1)$ and V_z^2 has $Q^{(1)} = (1, \tau_1, \tau_1^2 - a\tau_1^3)$, $Q^{(2)} = (\tau_2, 0, 1)$ where $\tau_1 = -t + at^2 - a^2 t^3$, $\tau_2 = -t$. The members of W_y have a common inflexion tangent whereas those of W_z do not. We note that for a form $\bar{F}$ to have both $\bar{V}_y^2$ and $\bar{V}_z^2$ projectively equivalent either to V_y^2 or to V_z^2 would contradict 7.3.

8. The nodal non-degenerate cases. A linear system of curves has no variable multiple point.[15] Hence, for every member of W_y to be of genus zero there must

[15] Bertini's theorem. [4] p. 25.

be a fixed double point (cusp). This double point counts four times as a member of V_y^2 but it is only three (linear) conditions on a cubic to have a fixed double point. For a cubic to contain the remaining two points (or branches) of V_y^2 is at most two further linear conditions. Hence, there is a $10 - 5 = 5$ parameter family of cubics on V_y^2 of which W_y is some 4 parameter subfamily.

We see readily that there are just five projective types of nodal cases:

N_1 : node plus two outside points.

N_2 : node with one fixed tangent plus an outside point.

N_3 : node with both tangents fixed.

N_4 : node with one tangent and a further fixed direction along that branch.

N_5 : cusp (the tangent must then be fixed).

Let t and u be indeterminants over K and set $R = K(u)[t] \bmod t^2$.

8.1 THEOREM. *If P and P' are images under T_1, then P is a node of W_y if and only if there exist polynomials linear in t and in u such that*

$$F(t, x_{\alpha\beta})P_\alpha P'_\beta x_{\alpha\beta} \equiv 0 \bmod t^2 \text{ identically in } u; \tag{8.2}$$

furthermore, if P is a node in W_y, then P' is a node in W_z.

PROOF: We suppose y and z coordinates so chosen that $P = (1, 0, 0)$ and $P' = (1, 0, 0)$. If P is a node of W_y, then any line $\rho y_1 = 1 = P_1$, $\rho y_2 = t = P_2$, $\rho y_3 = ut = P_3$, through P must meet the curves of W_y twice at P. Or stated analytically, if W_y is $G(\lambda, y) = \lambda_\delta f_\delta(y)$ then $G(\lambda, P_\alpha) \equiv 0 \bmod t^2$ identically in u.

Now since $\| y_{\beta\delta}(P_\alpha) \|$ is of rank 2 (in R), we proceed as in the proof of theorem 4.3 to obtain solutions $z_\beta = P'_\beta(t)$ of the equations 8.2 such that $P'_1 = 1$, $P'_2 = tf_2(u)$, $P'_3 = tf_3(u)$ where the f_β are rational functions of u with coefficients in K. A short argument (which we omit) shows that the form is degenerate unless f_2 and f_3 are linearly independent first degree polynomials in u, say $f_\beta(u) = c_\beta u + d_\beta$, $\beta = 1, 2$. The polynomials P, P' satisfy the conditions of the theorem. Further, since $\rho z_\beta = P'_\beta$ is by suitable choice of u, any line through P', 8.2 is symmetric in y and z so that P' is a node of W. This completes the proof of the necessity of 8.2. The sufficiency is obvious.

The cases $N_1, \cdots, N_4$ are sufficiently alike to justify our treating only one, N_1, in detail and merely listing a representative of each class in the other cases. Suppose then that V_y^2 consists of the points (1, 0, 0), (0, 1, 0), (0, 0, 1) of which the first is a node and let T_1 map these points into $z = (1, 0, 0), (0, 1, 0), (0, 0, 1)$ respectively. Now apply 8.2 and 4.3. We obtain

$$\| x_{\alpha\beta} \| = \begin{Vmatrix} 0 & x_{12} & x_{13} \\ d_2 x_{12} + d_3 x_{13} & 0 & x_{23} \\ c_2 x_{12} + c_3 x_{13} & x_{32} & 0 \end{Vmatrix}. \tag{8.4}$$

Let $x_{12} = x_2$, $x_{13} = x_3$, $x_{23} = x_1$, $x_{32} = x_4$. Then V_x^2 is

$$x_1 x_2 (c_2 x_2 + c_3 x_3) + x_3 x_4 (d_2 x_2 + d_3 x_3) = 0.$$

V_x^2 contains a line, l_n, of nodes, with the equations $x_2 = x_3 = 0$. The members of the pencil of planes $x_2 = \alpha x_3$ cut V_x^2 twice in l_n and in a further line

$$l_\alpha : x_1\alpha(\alpha c_2 + c_3) + x_4(\alpha d_2 + d_3) = 0, \quad x_2 = \alpha x_3 .$$

l_n and l_α meet at $(-\alpha d_2 - d_3, 0, 0, \alpha^2 c_2 + \alpha c_3)$. The point $(x_1, 0, 0, x_4)$ is on two distinct lines, l_α, l_{α^1}, unless the roots α, α^1 of

$$x_1(\alpha^2 c_2 + \alpha c_3) + x_4(\alpha d_2 + d_3) = 0$$

coincide. This happens if

$$(8.5) \qquad (c_3 x_1 + d_2 x_4)^2 - 4x_1 x_4 d_3 c_2 = 0.$$

The discriminant of 8.5 is

$$16 c_2 d_3 \begin{vmatrix} c_2 & c_3 \\ d_2 & d_3 \end{vmatrix}$$

which is zero only if the case is degenerate. Hence, 8.5 defines two distinct points, Q_1 and Q_2, on l_n. The points $R_1 = (1, 0, 0, 0)$ and $R_2 = (0, 0, 0, 1)$ are the only points of V_x^1.

8.6 Theorem. *Two forms F and $\bar{F}$ belonging to case N_1 are g-equivalent if and only if the unordered pair of points Q_1, Q_2 are projectively equivalent to the unordered pair $\bar{Q}_1$, $\bar{Q}_2$ under a projectivity which sends V_x^1 into $\bar{V}_x^1$.*

Proof: The "only if" is obvious since the points Q and R are defined purely in terms of V_x^2 and V_x^1. The "if" follows by normalization of 8.4 to a form containing a single parameter and then computation of the cross ratio of the four points in terms of this parameter. Subcases may arise in which one or both of the Q's may lie on V_x^1.

N_2 gives the single class represented by

$$\| x_{\alpha\beta} \| = \begin{Vmatrix} 0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \\ x_1 + x_2 & x_4 & 0 \end{Vmatrix}.$$

N_3 gives the single class represented by

$$\| x_{\alpha\beta} \| = \begin{Vmatrix} 0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \\ x_2 & x_4 & x_1 \end{Vmatrix}.$$

N_4 gives a family of g-classes with parameter $a \neq 0, -1$

$$\| x_{\alpha\beta} \| = \begin{Vmatrix} 0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \\ x_2 & a x_3 & x_4 \end{Vmatrix}.$$

Two forms F, F' in case N_4 are g-equivalent if and only if $a + 1/a = a' + 1/a'$.

The argument for the cusp case is only slightly different. Necessary and sufficient conditions for W_y to be cuspidal can be written in the form of 4.3 and

8.2 utilizing Puiseux expansions. There is just one class under N_5, represented by

$$\|x_{\alpha\beta}\| = \left\| \begin{matrix} 0 & x_1 & x_2 \\ -x_1 & x_2 & x_3 \\ -x_2 & -x_3 & x_4 \end{matrix} \right\|.$$

In every nodal case the principal manifold contains a line of nodes, and it is also a cone in cases N_3 and N_5.

9. The degenerate cases. Consideration of the manifolds V_y^1, V_z^1, V_x^1, V_y^2, V_z^2, V_x^2, and the relations between them defined by the transformations T_1, T_2, T_3 is sufficient to distinguish between the classes of degenerate forms in all save one set of classes, whose members are identical save for different values of a parameter and the distinctness of these classes follows from [14], p. 687; case 4.

The classes are first divided into sets of classes according to the nature of V_y^1 and V_z^1. Then among the classes possessing equivalent V_y^1 and equivalent V_z^1 we make a second subdivision according to V_x^1 or V_x^2 or both, and if this is still not sufficient to complete the classification we consider V_y^2 and V_z^2. In the appendix is a table which contains the matrix $\| y_{\beta\delta} \|$ for a representative of each degenerate class (3, 3, 4) and sufficient information to distinguish between the different classes. The proof of the completeness of the table is a normalization process along the lines of the similar proofs in [13] and [14] (i.e. successive discarding of forms equivalent to one retained). Because of its similarity to those and its length we shall omit the completeness proof.

We note that the terminology "degenerate cases" is justified in that V_x^2 is a reducible cubic in all of the degenerate classes except $g11$ and $g12$ in which it is a ruled cubic with a nodal line. Furthermore, it is also true that no manifolds can be V_x^2 for both a degenerate and a linear non-degenerate class, so that it would have been possible to have defined "degenerate class" in terms of V_x^2 alone, proper account being taken of classes $g11$ and $g12$.

As a partial summary of the results listed in the table of classes we give the following theorem which solves problem I for the degenerate cases.

9.1 THEOREM. *A necessary and sufficient condition for the equivalence of two degenerate trilinear forms F and $\bar{F}$ with $(p, q, r) = (3, 3, 4)$ is that V_t^1 and V_t^2 be simultaneously projectively equivalent to $\bar{V}_t^1$ and $\bar{V}_t^2$ for $t = x, y, z$ in turn, except in the following cases:*

a) *If V_x^1, V_y^1, V_z^1 consist of* 2, 1, 1 *points respectively, V_x^2 of three independent planes, and T_3, T_2 map V_y^1, V_z^1 into π_1, π_2 respectively,*[16] *and the same for $\bar{F}$, then $F \sim \bar{F}$ if and only if* i) $\pi_1 = \pi_2$ *and* $\bar{\pi}_1 = \bar{\pi}_2$ *or* ii) $\pi_1 \neq \pi_2$ *and* $\bar{\pi}_1 \neq \bar{\pi}_2$.

b) *If V_x^1, V_y^1, V_z^1 consist of* 1, 0, 0 *points respectively and V_y^2 is a line; and the same for $\bar{F}$, then $F \sim \bar{F}$ if and only if $a + 1/a = \bar{a} + 1/\bar{a}$ where a, $\bar{a}$ are the parameters of F and $\bar{F}$ in the canonical form given for the classes $g_1 2$ in the table.*

[16] π_1, π_2 being planes of V_x^2.

Of the 53 degenerate cases (the infinite family of g-classes discussed in 9.1b are counted as one case, g_12, but each other case includes only a single g-class) consideration of V_y^1, V_z^1, V_z^2, and V_x^1 is sufficient to isolate 37 and to divide the remaining 16 into 8 pairs. In four of these pairs: e_18, e_38; e_58, e_68; e_16, e_56; f_12, f_72 the distinctness is a consequence of different projective situations of V_x^1 on V_x^2. The distinction between e_36 and e_66 is given in 9.1a. In the remaining three pairs; f_32, f_42; f_21; f_31; g_22, g_32 consideration of V_y^2 and V_z^2 is sufficient, and these are the only degenerate cases which require consideration of V_y^2 and V_z^2.

10. Trilinear forms with $3 \leqq p < q \leqq r$ or $3 < p = q \leqq r$. Although not formulated in the language of trilinear forms the work of Room[17] on the freedom of a projectively generated (determinantal) manifold can be rephrased to give the following theorems:

10.1 Theorem. *In general, two forms F and G with $3 \leqq p < q \leqq r$ are projectively equivalent if and only if their principal manifolds are projectively equivalent.*

10.2 Theorem. *In general, two forms F and G with $3 < p = q \leqq r$ are g-equivalent if and only if their principal manifolds are projectively equivalent.*

In a sense 10.1 and 10.2 solve problem I for all cases (p, q, r) not treated in [14] and the earlier sections of this paper. However, the qualifying "in general" leaves much to be desired and requires some clarification.

There are at least two current interpretations of "in general." The modern algebraist means by "general" that the coefficients are independent indeterminants over the base field. The traditional algebraic geometer means by "general" that the coefficients are given elements of the base field but are not in any special relation to each other. In this discussion we shall denote the first interpretation by "generic" and the second by "non-special." The above theorems refer to the second interpretation.

Two completely generic trilinear forms with the same p, q, r are isomorphic but the concept of projective equivalence is meaningless (unless the indeterminants of the two forms are related). On the other hand, projective equivalence always has meaning for non-special forms, but the concept non-special is rather vague and can be made definite for objects of a given category only after the relevant properties of these objects are known.

We can illustrate this by discussing several possible definitions of non-special forms (3, 3, 4). We consider first

10.3 Definition. A form $F(3, 3, 4)$ is said to be *non-special* if

a) V_y^1 is empty.
b) V_y^2 consists of 6 distinct points not on a conic and no 3 on a line.
c) V_y^2 is not projectively equivalent to V_z^2.

With this definition of non-special the following theorem (cf. theorem 6.6) is valid.

[17] [3] Chapter VII, especially p. 124, or [10].

10.4 Theorem. *In general, $F \sim \bar{F}$ if and only if V_y^2, V_z^2 are projectively equivalent to $\bar{V}_y^2$, $\bar{V}_z^2$, respectively.*

However, 10.4 remains true (cf. 7.2) with a much weaker definition of "in general." We may delete 10.3c and replace 10.3b by

(10.3b′) V_y^2 is zero dimensional of order 6 and W_y is non-nodal.

But this modified definition could hardly have been made without the complete solutions of problems I and II for forms (3, 3, 4) at hand.

The "only if" of 10.4 is valid for any definition of "in general." The "if" follows for any non-degenerate case in which V_y^2 defines a unique W_y. This is the real reason for distinguishing between the non-nodal and nodal cases.

The situation for (3, 3, 4) suggests that the non-special need not be of the strongest possible type and leads to the question: Just how special can a trilinear form (p, q, r) be and still be non-special with respect to theorems 10.1 or 10.2? It is certain that the concept "non-special" must include some generalization of the concepts "non-degenerate" and "nodal," and it is possible that no further requirements need be made.

Theorems 10.1 and 10.2 give no clue to the solution of problem II. For forms with $p = q$, Room[18] has a theorem which gives necessary and sufficient conditions that a hypersurface can be a principal manifold, but these conditions being inductive in nature are not completely satisfactory from our point of view. However, in any systematic attempt at solving problem II one would do well to take account of the numerous examples and theorems contained in Room's work.

Appendix. The cases in the table of g-classes are labeled first with a small letter describing V_y^1 and V_z^1 and then a number giving the projective nature of V_z^2. When there are several cases with the same letter and number, subscripts are attached to the letter. The capital letters after the comma describe V_z^1 and are not a part of the label. The key to the letters and numbers used follows:

V_y^1 and V_z^1 are

- *a* both lines
- *b* one a line and one a point
- *c* both points
- *d* one two points and one a single point
- *e* both a single point
- *f* one a single point and one empty
- *g* both empty

V_z^2 is

1. quadric and plane meeting in a conic
2. quadric and plane meeting in two lines
3. cone and plane meeting in a conic
4. cone and plane meeting in two distinct lines
5. cone and plane meeting in a single line

[18] [3] p. 65.

6 three independent planes
7 three coaxial planes
8 two planes one double
9 triple plane
10 whole space
11 ruled cubic with nodal line of first kind[19]
12 ruled cubic with nodal line of second kind[19]

In the description of V_x^1

P single point
P_i i points, $i > 1$
L one line
$\bar{L}_2$ two intersecting lines
L_2 two skew lines
C conic
Q plane
E empty

(In combinations such as LP_i read "line plus i points.")

TABLE OF DEGENERATE g-CLASSES

$a_1 10, \bar{L}_2 \quad \left\| \begin{matrix} y_1 & y_2 & 0 & 0 \\ 0 & 0 & y_3 & 0 \\ 0 & 0 & 0 & y_3 \end{matrix} \right\| \qquad a_2 10, L_2 \quad \left\| \begin{matrix} y_1 & y_2 & y_3 & 0 \\ 0 & y_3 & 0 & 0 \\ 0 & 0 & 0 & y_3 \end{matrix} \right\|$

$b8, Q \quad \left\| \begin{matrix} y_1 & 0 & 0 & y_3 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_3 & 0 & 0 \end{matrix} \right\| \qquad c6, LP_2 \quad \left\| \begin{matrix} y_1 & 0 & y_3 & 0 \\ 0 & y_2 & 0 & 0 \\ 0 & 0 & 0 & y_3 \end{matrix} \right\|$

$d_1 6, LP_2 \quad \left\| \begin{matrix} y_1 & 0 & y_3 & 0 \\ 0 & y_2 & y_3 & 0 \\ 0 & 0 & 0 & y_3 \end{matrix} \right\| \qquad d_2 6, LP_1 \quad \left\| \begin{matrix} y_1 & 0 & y_3 & 0 \\ y_3 & y_2 & 0 & 0 \\ 0 & 0 & 0 & y_3 \end{matrix} \right\|$

$e9, L_2 \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ y_2 & 0 & 0 & y_3 \\ y_1 & y_2 & y_3 & 0 \end{matrix} \right\| \qquad e_1 8, C \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ y_1 & 0 & y_2 & 0 \end{matrix} \right\|$

$e_2 8, L_2 \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ y_1 & y_3 & y_2 & 0 \end{matrix} \right\| \qquad e_3 8, C \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_1 & 0 & y_2 & y_3 \end{matrix} \right\|$

$e_4 8, LP \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_1 & 0 & y_2 \end{matrix} \right\| \qquad e_5 8, L \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ y_2 & y_3 & 0 & 0 \\ 0 & y_1 & y_2 & y_3 \end{matrix} \right\|$

$e_6 8, L \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_1 & y_2 & y_3 \end{matrix} \right\| \qquad e_1 7, P_2 \quad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_1 & y_1 & 0 & y_2 \end{matrix} \right\|$

---

[19] See [1], p. 61.

$$e_27,\ P \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_1 & y_1 & y_2 & y_3 \end{matrix} \right\|$$

$$e_26,\ P_4 \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ 0 & y_3 & y_1 & y_2 \end{matrix} \right\|$$

$$e_46,\ LP \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ y_2 & 0 & y_1 & 0 \end{matrix} \right\|$$

$$e_66,\ P_2 \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ y_2 & y_3 & y_1 & 0 \end{matrix} \right\|$$

$$e2,\ L_2 \qquad \left\| \begin{matrix} y_1 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_3 & 0 & 0 & y_2 \end{matrix} \right\|$$

$$f_15,\ P_2 \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_2 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_35,\ P \qquad \left\| \begin{matrix} y_1 & y_2 & 0 & y_3 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_24,\ P_3 \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_2 + y_3 \\ y_2 & 0 & y_3 & 0 \\ y_3 & y_2 & 0 & 0 \end{matrix} \right\|$$

$$f_13,\ P_2 \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_2 \\ 0 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_33,\ P_4 \qquad \left\| \begin{matrix} y_1 & y_2 + y_3 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_22,\ P_3 \qquad \left\| \begin{matrix} y_1 & y_2 & y_3 & 0 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_42,\ L \qquad \left\| \begin{matrix} y_1 & 0 & y_3 & 0 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_62,\ LP \qquad \left\| \begin{matrix} y_1 & 0 & y_2 & 0 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$e_16,\ P_3 \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ 0 & 0 & y_1 & y_2 \end{matrix} \right\|$$

$$e_36,\ P_2 \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_2 & 0 & y_1 & y_3 \end{matrix} \right\|$$

$$e_56,\ P_3 \qquad \left\| \begin{matrix} y_3 & 0 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ y_2 & 0 & y_1 & y_3 \end{matrix} \right\|$$

$$e3,\ CP \qquad \left\| \begin{matrix} y_1 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ 0 & 0 & y_2 & y_3 \end{matrix} \right\|$$

$$e1,\ C \qquad \left\| \begin{matrix} y_1 & 0 & 0 & 0 \\ 0 & y_2 & y_3 & 0 \\ y_2 & 0 & y_2 & y_3 \end{matrix} \right\|$$

$$f_25,\ L \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_3 \\ y_2 & 0 & y_3 & 0 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_14,\ P_2 \qquad \left\| \begin{matrix} y_1 & y_3 & 0 & y_2 \\ y_2 & 0 & y_3 & 0 \\ y_3 & y_2 & 0 & 0 \end{matrix} \right\|$$

$$f_34,\ LP \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_2 \\ y_2 & 0 & y_3 & 0 \\ y_3 & y_2 & 0 & 0 \end{matrix} \right\|$$

$$f_23,\ P_3 \qquad \left\| \begin{matrix} y_1 & y_2 & 0 & 0 \\ 0 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_12,\ P_2 \qquad \left\| \begin{matrix} y_1 & y_2 & 0 & 0 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_32,\ L \qquad \left\| \begin{matrix} y_1 & 0 & 0 & y_2 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_52,\ P \qquad \left\| \begin{matrix} y_1 & 0 & y_3 & y_2 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_72,\ P_2 \qquad \left\| \begin{matrix} y_1 & 0 & y_2 & y_2 \\ y_2 & 0 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{matrix} \right\|$$

$$f_1 1, P \quad \begin{Vmatrix} y_1 & 0 & 0 & y_2 \\ y_2 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{Vmatrix} \qquad f_2 1, P_2 \quad \begin{Vmatrix} y_1 & y_2 & 0 & 0 \\ y_2 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{Vmatrix}$$

$$f_3 1, P_2 \quad \begin{Vmatrix} y_1 & y_2 & y_2 & 0 \\ y_2 & y_2 & 0 & y_3 \\ 0 & y_3 & y_2 & 0 \end{Vmatrix} \qquad g12, L \quad \begin{Vmatrix} y_1 & y_2 & y_3 & 0 \\ 0 & 0 & y_2 & y_1 \\ 0 & 0 & y_1 & y_3 \end{Vmatrix}$$

$$g11, L \quad \begin{Vmatrix} y_1 & y_2 & 0 & 0 \\ 0 & 0 & y_2 & y_3 \\ 0 & 0 & y_3 & y_1 \end{Vmatrix} \qquad g8, P \quad \begin{Vmatrix} 0 & y_3 & -y_2 & y_1 \\ -y_3 & 0 & y_1 & 0 \\ y_2 & -y_1 & 0 & 0 \end{Vmatrix}$$

$$g6, P \quad \begin{Vmatrix} 0 & y_3 & -y_2 & 0 \\ -y_3 & 0 & y_1 & y_1 \\ y_2 & -y_1 & 0 & 0 \end{Vmatrix} \qquad g3, P_2 \quad \begin{Vmatrix} y_2 & y_3 & 0 & 0 \\ 0 & 0 & y_2 & y_1 \\ 0 & y_1 & y_3 & 0 \end{Vmatrix}$$

$$g_1 2, P \quad \begin{Vmatrix} y_2 & y_3 & 0 & y_1 \\ 0 & ay_1 & y_2 & 0 \\ y_1 & 0 & y_3 & 0 \end{Vmatrix} \qquad g_2 2, P_2 \quad \begin{Vmatrix} y_2 & y_3 & y_1 & 0 \\ 0 & 0 & y_2 & y_1 \\ y_1 & 0 & y_3 & 0 \end{Vmatrix}$$

$$g_3 2, P_2 \quad \begin{Vmatrix} y_2 & y_3 & 0 & 0 \\ 0 & 0 & y_2 & y_1 \\ y_1 & 0 & y_3 + y_1 & 0 \end{Vmatrix} \qquad g_1 1, E \quad \begin{Vmatrix} 0 & y_3 & -y_2 & y_1 \\ -y_3 & 0 & y_1 & 0 \\ y_2 & -y_1 & 0 & -y_2 \end{Vmatrix}$$

$$g_2 1, P_2 \quad \begin{Vmatrix} y_2 & y_3 & y_1 & 0 \\ 0 & 0 & y_2 & y_1 \\ 0 & y_1 & y_3 & 0 \end{Vmatrix}$$

UNIVERSITY OF MICHIGAN.

BIBLIOGRAPHY

(1) A. B. Bassett, *A Treatise on the Geometry of Surfaces*, Cambridge, 1910.

(2) Julian L. Coolidge, *Algebraic Plane Curves*, Oxford, 1931.

(3) T. G. Room, *The Geometry of Determinantal Loci*, Cambridge, 1938.

(4) Severi-Löffler, *Vorlesungen über Algebraische Geometrie*, Teubner, 1921.

(5) B. L. Van der Waerden, *Moderne Algebra*, Springer, 1931.

(6) Josephine Chanler, *The Invariant Theory of the Ternary Trilinear Form*, Duke Mathematical Journal, vol. 5 (1939), pp. 552–566.

(7) A. B. Coble, *Point Sets and Allied Cremona Groups, Part I*, Transactions of the American Mathematical Society, vol. 16 (1915), pp. 155–198.

(8) L. E. Dickson, *Determination of all General Homogeneous Polynomials Expressible as Determinants with Linear Elements*, Transactions of the American Mathematical Society, vol. 22 (1921), pp. 167–179.

(9) L. E. Dickson, *Singular Case of Pairs of Bilinear, Quadratic, or Hermitian Forms*, Transactions of the American Mathematical Society, vol. 29 (1927), pp. 239–253.

(10) T. G. Room, *Freedoms of Determinantal Manifolds*, Proceedings of the London Mathematical Society, (2) vol. 36 (1934), pp. 1–28.

(11) Virgil Snyder, *Room on Determinantal Loci*, Bulletin of the American Mathematical Society, vol. 45 (1939), pp. 499–501.

(12) R. M. Thrall, *Apolarity of Trilinear Forms and Pencils of Bilinear Forms*, Bulletin of the American Mathematical Society, vol. 44 (1938), pp. 678–683.

(13) R. M. Thrall, *Metabelian Groups and Trilinear Forms*, American Journal of Mathematics, vol. 60 (1937), pp. 383–415.

(14) R. M. Thrall and Josephine Chanler, *Ternary Trilinear Forms in the Field of Complex Numbers*, Duke Mathematical Journal, vol. 4 (1938), pp. 678–690.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

HOMOMORPHISM OF GROUPS

BY J. H. M. WEDDERBURN

(Received October 18, 1940)

As usually given homomorphism between groups is a many-one relation but it will be shown below that the treatment of the many-many relation is just as simple and in some respects clearer because of the symmetry.

Let G and H be two groups each divided into mutually exclusive sets

$$G = G_1 + G_2 + \cdots = \sum G_p, \qquad H = H_1 + H_2 + \cdots = \sum H_p$$
$$G_p \frown G_q = 0, \qquad H_p \frown H_q = 0 \qquad (p \neq q)$$

where the notation is not to be taken to mean that the sets are denumerable; the elements of a set G_p will be denoted by $g_p, g'_p, \cdots$. Then G is said to be homomorphic to H, $G \sim H$, if

$$(1) \qquad g_pg_q = g_r \supset H_pH_q \leqq H_r.$$

If $g'_pg'_q = g_s$, then $H_pH_q \leqq H_s$; but $H_r \frown H_s = 0$ if $r \neq s$; hence $r = s$, that is, (1) implies

$$(2) \qquad G_pG_q \leqq G_r.$$

Again, if $h_ah_b = h_c$, let $g_ag_b = g_d$; then $H_aH_b \leqq H_d$. But $H_aH_b \geqq h_c < H_c$; therefore $H_c \frown H_d \neq 0$ so that $d = c$ and, since $g_d < G_d$ no matter what elements g_a and g_b are in G_a and G_b, we have

$$h_ah_b = h_c \supset G_aG_b \leqq G_c$$

that is, the relation of homomorphism is reflexive.

Suppose now that the identity g_1 is in G_1; then

$$g_1g_p = g_p = g_pg_1 \supset G_1G_p = G_p = G_pG_1$$

for all p, and therefore

$$(3) \qquad H_1H_p \leqq H_p, \qquad H_pH_1 \leqq H_p.$$

If the identity of H lies in H_a, it follows similarly that

$$H_aH_p = H_p = H_pH_a$$

and in particular

$$H_aH_1 = H_1 = H_1H_a.$$

But from (3) $H_aH_1 \leqq H_a$ and hence $a = 1$, so that the identity of H lies in H_1.

Since G is a group, any g'_1 has an inverse, say g_p, such that

$$g'_1g_p = g_1 = g_pg'_1.$$

Hence $G_1G_p \leqq G_1$: but $G_1G_p = G_p$ and hence $G_p = G_1$, that is, G_1 is a group. Since the relation is reflexive, H_1 is also a group. Let g_p be any element of G_p $(p \neq 1)$ and let $g_p^{-1} = g_q < G_q$; then $G_pG_q \frown G_1 \geqq g_1$ and therefore from (2) $G_pG_q \leqq G_1$. But $G_pg_q \leqq G_pG_q$ so that $G_pg_q \leqq G_1$. But $G_pg_q \leqq G_pG_q$ so that $G_pg_q \leqq G_1$; hence

$$G_p = G_pg_qg_p \leqq G_1g_p \leqq G_p$$

so that

$$(4) \qquad G_p = G_1g_p$$

for every element g_p in G_p. Similarly $H_p = H_1h_p$ and $G_p = g_pG_1$, $H_p = h_pH_1$, so that G_1 and H_1 are invariant in G and H. The final result can now be stated.

THEOREM. *If $G \sim H$, then also $H \sim G$. If G_1 contains the identity of G, then $H_1 \sim G_1$ contains the identity of H. Further G_1 and H_1 are invariant subgroups of G and H, respectively, and $G/G_1 \simeq H/H_1$.*

Let G_1 be a subgroup of G minimal with respect to the property that there is a homomorphism with H given by $G/G_1 \simeq H/H_1$; and let G_2 be a second such subgroup. If we set

$$G = \sum G_1g_i = \sum G_2g_i', \qquad H = \sum H_1h_i = \sum H_2h_i'$$

then $g_i \sim h_i$ in the first homomorphism, and $g_i' \sim h_i'$ in the second. Let

$$B = G_1 \frown G_2, \qquad G_1 = \sum B\gamma_i, \qquad G_2 = \sum B\gamma_i'$$

$$C = H_1 \frown H_2, \qquad H_1 = \sum C\eta_i, \qquad H_2 = \sum C\eta_i';$$

then $G_1G_2 = \sum B\gamma_i\gamma_j'$, and

$$G = \sum G_1G_2\gamma_k'' = \sum B\gamma_i\gamma_j'\gamma_k''$$

$$H = \sum H_1H_2\eta_k'' = \sum C\eta_i\eta_j'\eta_k''$$

which gives a homomorphism between G and H by means of B and C. But B is a subgroup of G_1 which is minimal and hence $B = G_1$ and so $G_2 = G_1$. The minimal subgroup G_1 is therefore unique.

PRINCETON, NEW JERSEY.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

AN ANALOGUE TO MINKOWSKI'S GEOMETRY OF NUMBERS IN A FIELD OF SERIES

BY KURT MAHLER

(Received November 16, 1939)

Minkowski, in his "Geometrie der Zahlen" (Leipzig 1910), studied properties of a convex body in a space R_n of n dimensions with respect to the set of all lattice points. Let $F(X) = F(x_1, \cdots, x_n)$ be a distance function, i.e. a function satisfying the conditions

$$F(0) = 0, F(X) > 0 \text{ if } X \neq 0;$$

$$F(tX) = |t| F(X) \text{ for all real } t;$$

$$F(X - Y) \leqq F(X) + F(Y).$$

The inequality $F(X) \leqq 1$ defines a convex body in R_n which has its centre at the origin $X = 0$. Suppose that this body has the volume V. The well known result of Minkowski asserts that *if* $V \geqq 2^n$, *then the body contains at least one* (*and so at least two*) *lattice points different from* 0. This theorem is contained in the following deeper result of Minkowski (G.d.Z. §§50–53): "*There are n independent lattice points $X^{(1)}, X^{(2)}, \cdots, X^{(n)}$ in R_n with the following properties*: (1) $F(X^{(1)}) = \sigma^{(1)}$ *is the minimum of $F(X)$ in all lattice points $X \neq 0$, and for $k \geqq 2$, $F(X^{(k)}) = \sigma^{(k)}$ is the minimum of $F(X)$ in all lattice points X which are independent of $X^{(1)}, \cdots, X^{(k-1)}$.* (2) *The determinant D of the points $X^{(1)}, \cdots, X^{(n)}$ satisfies the inequalities*

$$1 \leqq |D| \leqq n!.$$

(3) *The numbers $\sigma^{(k)}$ depend only on $F(X)$ and not on the special choice of the lattice points $X^{(k)}$, and they satisfy the inequalities*

$$0 < \sigma^{(1)} \leqq \sigma^{(2)} \leqq \cdots \leqq \sigma^{(n)}, \qquad \frac{2^n}{n!} \leqq V\sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)} \leqq 2^n."$$

(A new simple proof for the last part of this theorem was given by *H. Davenport*, Quart. Journ. Math. (Oxford Ser.), Vol. 10 (1939), 119–121).

From Minkowski's theorem, properties of general classes of convex bodies can be obtained. For instance, there is a *convex body* $G(Y) \leqq 1$ *polar to* $F(X) \leqq 1$, and to this body correspond by the theorem n minima $\tau^{(1)}, \tau^{(2)}, \cdots, \tau^{(n)}$. I have proved (Časopis 68 (1939), 93–102), that *these minima are related to the σ's by the inequalities*

$$1 \leqq \sigma^{(h)}\tau^{(n-h+1)} \leqq (n!)^2 \qquad (h = 1, 2, \cdots, n).$$

From this result, applications to inhomogeneous Diophantine inequalities can be made, and in particular, generalizations of *Kronecker's* theorem can be obtained.

The present paper does *not* deal with ordinary convex bodies in a real space. The n-dimensional space P_n with which we shall be concerned has its coordinates in a field $\mathfrak{K}$ with a non-*Archimedean* valuation $|x|$; a distance function is any function satisfying

$$F(0) = 0, F(X) > 0 \text{ if } X \neq 0,$$

$$F(tX) = |t| F(X) \text{ for all } t \text{ in } \mathfrak{K},$$

$$F(X - Y) \leqq \max (F(X), F(Y)).$$

The inequality $F(X) \leqq \tau$ then defines the convex body $C(\tau)$, if $\tau > 0$. We show that every convex body is bounded, and that it has properties similar to a parallelepiped in real space.

In particular, let $\mathfrak{K}$ be the field of all Laurent series

$$x = \alpha_f z^f + \alpha_{f-1} z^{f-1} + \alpha_{f-2} z^{f-2} + \cdots$$

with coefficients in an arbitrary field $\mathfrak{k}$; the valuation $|x|$ is defined as $|0| = 0$, and $|x| = e^f$ if $\alpha_f \neq 0$. Further let Λ_n be the modul of all points in P_n, the coordinates of which are polynomials in z with coefficients in $\mathfrak{k}$; these points we call *lattice points*. We consider only distance functions $F(X)$ which for all $X \neq 0$ in P_n are always as integral power of e. We shall define a certain positive constant V as the volume of $C(1)$; this constant is invariant under all linear transformations of P_n with determinant 1, and the volume of $C(1)$ and that of its polar reciprocal body $C'(1)$ have the product 1. In analogy to Minkowski's theorem, the following theorem holds: "*There are n independent lattice points* $X^{(1)}, \cdots, X^{(n)}$ *in* P_n *with the following properties*: 1) $F(X^{(1)})$ *is the minimum of* $F(X)$ *in all lattice points* $X \neq 0$, *and for* $k \geqq 2$, $F(X^{(k)})$ *is the minimum of* $F(X)$ *in all lattice points X which are independent of* $X^{(1)}, \cdots, X^{(k-1)}$. 2) *The determinant of the points* $X^{(1)}, \cdots, X^{(n)}$ *is* 1. 3) *The numbers* $F(X^{(k)}) = \sigma^{(k)}$, *which depend only on* $F(X)$ *and not on the special choice of the lattice points X, satisfy the formulae*

$$0 < \sigma^{(1)} \leqq \sigma^{(2)} \leqq \cdots \leqq \sigma^{(n)}, \qquad \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)} = \frac{1}{V}.\text{"}$$

Further, we have similar minima $\tau^{(1)}, \cdots, \tau^{(n)}$ for the distance function $G(Y)$ which defines the polar body $C'(1)$; these are related with the σ's by the equations

$$\sigma^{(h)}\tau^{(n-h+1)} = 1 \qquad (h = 1, 2, \cdots, n).$$

These two results can be used to study special Diophantine problems in P_n; a few of them are considered as examples. All the proofs in this paper are based on the methods of Minkowski, and in one final paragraph I make use of ideas of C. L. Siegel.

I. Convex Domains in Non-Archimedean Spaces

1. Notation. In this chapter, we denote by

$\mathfrak{R}$ *an arbitrary field,*

$|x|$ *a non-Archimedean valuation of the elements x of $\mathfrak{R}$,*[1]

$\mathfrak{K}$ *the perfect extension of $\mathfrak{R}$ with respect to this valuation,*

P_n *the n-dimensional space of all points or vectors*

$$X = (x_1, \cdots, x_n),$$

where the coordinates $x_1, \cdots, x_n$ lie in $\mathfrak{K}$,

$|X|$ *the length of the vector X, viz.*

$$|X| = \max(|x_1|, \cdots, |x_n|).$$

We apply the usual notation for vectors in P_n; thus if

$$X = (x_1, \cdots, x_n) \quad \text{and} \quad Y = (y_1, \cdots, y_n),$$

and a belongs to $\mathfrak{K}$, then we write

$$X \mp Y = (x_1 \mp y_1, \cdots, x_n \mp y_n),$$

$$aX = (ax_1, \cdots, ax_n),$$

$$XY = \sum_{h=1}^{n} x_h y_h.$$

For instance, the length $|X|$ of X has the properties:

(1) $\quad |X| \geqq 0$, with equality if and only if $X = (0, \cdots, 0) = 0$;

(2) $\quad |aX| = |a||X|$, if a is any element of $\mathfrak{K}$;

(3) $\quad |X \mp Y| \leqq \max(|X|, |Y|)$;

(4) $\quad |XY| \leqq |X||Y|$.

If $\mathfrak{D}$ is any sub-ring of $\mathfrak{K}$, and $X^{(1)}, \cdots, X^{(r)}$ are vectors in P_n, then these are called $\mathfrak{D}$-dependent, or $\mathfrak{D}$-independent, according as there exist, or do not exist elements $a_1, \cdots, a_r$ of $\mathfrak{D}$ not all zero, such that

$$a_1X^{(1)} + \cdots + a_rX^{(r)} = 0.$$

A set of vectors of P_n is called a $\mathfrak{D}$-modul, if with X and Y it also contains $aX + bY$, where a and b are arbitrary elements of $\mathfrak{D}$; the modul has the dimen-

---

[1] This means that the function $|x|$ satisfies the conditions:

$$|0| = 0, \text{ but } |x| > 0 \text{ for } x \neq 0,$$

$$|xy| = |x||y|,$$

$$|x \mp y| \leqq \max(|x|, |y|).$$

sion m, if there are m, but not $m + 1$, $\mathfrak{D}$-independent elements in it. The dimension of a $\mathfrak{K}$-modul is at most n, while that of any other class of moduli need not be finite.

2. The distance function $F(X)$. A function $F(X)$ of the variable point X in P_n is called a general distance function, if it has the properties:

(A): $$F(X) \geqq 0;$$

(B): $$F(aX) = |a| F(X) \text{ for all } a \text{ in } \mathfrak{K}, \text{ hence } F(0) = 0;$$

(C): $$F(X \mp Y) \leqq \max (F(X), F(Y));$$

it is called a special distance function or simply a distance function, if instead of (A) it satisfies the stronger condition

(A'): $$F(X) > 0 \text{ for } X \neq 0.$$

If τ is a positive number, then the set $C(\tau)$ of all points X with

$$F(X) \leqq \tau$$

is called a convex set;[2] if $F(X)$ is a special distance function, then it is called a convex body. It is clear from the definition of $F(X)$ that a convex set $C(\tau)$ contains the origin 0, and that with X and Y also $aX + bY$ belong to it, if a and b are elements of $\mathfrak{K}$ such that $|a| \leqq 1$, $|b| \leqq 1$. Further, if

$$E^{(1)} = (1, 0, \cdots, 0), E^{(2)} = (0, 1, \cdots, 0), \cdots, E^{(n)} = (0, 0, \cdots, 1)$$

are the n unit vectors of the coordinate system, then

$$X = x_1 E^{(1)} + \cdots + x_n E^{(n)}, \quad \text{i.e.} \quad F(X) \leqq \max_{h=1,2,\cdots,n} (|x_h| F(E^{(h)})),$$

and therefore

(5) $$F(X) \leqq \Gamma |X|,$$

where Γ is the positive constant

$$\Gamma = \max_{h=1,2,\cdots,n} (F(E^{(h)})).$$

$C(\tau)$ contains therefore all points of the cube

$$|X| \leqq \frac{\tau}{\Gamma}.$$

We prove now that *for special distance functions there is a* second *positive constant* γ, *such that for all points in* P_n

(6) $$F(X) \geqq \gamma |X|.$$

[2] We consider only convex sets and bodies as defined; they are obviously symmetrical with respect to the origin.

PROOF: We assume that (6) is not true and show that this leads to a contradiction.

By hypothesis, there is an infinite sequence S of points

$$X^{(h)} = (x_1^{(h)}, \cdots, x_n^{(h)}) \neq 0 \qquad (h = 1, 2, 3, \ldots),$$

such that

$$\lim_{h\to\infty} \frac{F(X^{(h)})}{|X^{(h)}|} = 0.$$

Since

$$\frac{F(aX)}{|aX|} = \frac{F(X)}{|X|}$$

for all $a \neq 0$ in $\mathfrak{K}$, we may assume that for the elements of S

$$\lim_{h\to\infty} F(X^{(h)}) = 0, \qquad |X^{(h)}| = 1,$$

so that in particular the n real sequences

$$|x_k^{(1)}|, |x_k^{(2)}|, |x_k^{(2)}|, \cdots \qquad (k = 1, 2, \cdots, n)$$

are bounded.

Hence we can replace S by an infinite sub-sequence which we again call S: $X^{(1)}, X^{(2)}, X^{(3)}, \cdots$, such that the n real limits

$$(7) \qquad a_k = \lim_{h\to\infty} |x_k^{(h)}| \qquad (k = 1, 2, \cdots, n)$$

exist and satisfy the equation

$$\max_{k=1,2,\cdots,n} a_k = 1.$$

We call S a sequence of rank m, if exactly m of the limits $a_1, a_2, \cdots, a_n$ do not vanish; without loss of generality, these are the m first limits $a_1, a_2, \cdots, a_n$. Obviously $1 \leqq m \leqq n$.

If the rank $m = 1$, then for large h

$$|x_1^{(h)}| = 1, \quad \text{and} \quad \frac{X^{(h)}}{x_1^{(h)}} = \left(1, \frac{x_2^{(h)}}{x_1^{(h)}}, \cdots, \frac{x_n^{(h)}}{x_1^{(h)}}\right), = E^{(1)} + X^{*(h)}$$

say, where

$$\lim_{h\to\infty} |X^{*(h)}| = 0.$$

Hence by (5)

$$F(E^{(1)}) = F\left(\frac{X^{(h)}}{x_1^{(h)}} - X^{*(h)}\right) \leqq \max\left(\frac{F(X^{(h)})}{|x_1^{(h)}|}, F(X^{*(h)})\right)$$
$$\leqq \max\,(F(X^{(h)}), \Gamma\,|X^{*(h)}|),$$

and therefore for $h \to \infty$

$$0 \leqq F(E^{(1)}) \leqq 0, \text{ i.e. } F(E^{(1)}) = 0,$$

which is not true.

Hence the rank $m \geqq 2$. Put

$$X^{(g,h)} = \frac{X^{(g)}}{x_m^{(g)}} - \frac{X^{(h)}}{x_m^{(h)}} = (x_1^{(g,h)}, \cdots, x_m^{(g,h)}).$$

Then from (7) for large g, h

$$F(X^{(g,h)} \leqq \max\left(\frac{F(X^{(g)})}{|x_m^{(g)}|}, \frac{F(X^{(h)})}{|x_m^{(h)}|}\right) \leqq \frac{2}{a_m} \max(F(X^{(g)}), F(X^{(h)})),$$

and therefore

$$\lim_{\substack{g\to\infty\\h\to\infty}} F(X^{(g,h)}) = 0.$$

Two cases are now possible:

a: The limit

$$\lim_{\substack{g\to\infty\\h\to\infty}} | X^{(g,h)} | = \lim_{\substack{g\to\infty\\h\to\infty}} \max(| x_1^{(g,h)} |, \cdots, | x_n^{(g,h)} |)$$

exists and is zero. Hence the n limits in $\mathfrak{K}$

(8) $$x_k^* = \lim_{h\to\infty} \frac{x_k^{(h)}}{x_m^{(h)}} \qquad (k = 1, 2, \cdots, n)$$

all exist, and in particular

$$x_m^* = \lim_{h\to\infty} 1 = 1,$$

so that

$$X^* = (x_1^*, \cdots, x_n^*) \neq 0.$$

By the continuity of $F(X)$,[3]

$$F(X^*) = \lim_{h\to\infty} F\left(\frac{X^{(h)}}{x_m^{(h)}}\right) = \frac{1}{a_m} \lim_{h\to\infty} F(X^{(h)}) = 0,$$

which is not true.

b: The limit

$$\lim_{\substack{g\to\infty\\h\to\infty}} | X^{(g,h)} |$$

[3] If $\epsilon > 0$ is given, then there is a $\delta > 0$, such that $| F(X) - F(Y) | < \epsilon$ for $| X - Y | < \delta$, as follows easily from the properties (B), (C), and (5).

either does not exist, or exists and is different from zero. That implies that at least one of the limits (8) does not exist. Now obviously

$$\lim_{\substack{g\to\infty\\h\to\infty}} |x_k^{(g,h)}| = 0 \qquad (k = m, m+1, \cdots, n),$$

since for large g, h

$$x_m^{(g,h)} = 0; \qquad |x_k^{(g,h)}| = \left|\frac{x_k^{(g)}}{x_m^{(g)}} - \frac{x_k^{(h)}}{x_m^{(h)}}\right| \leqq \frac{2}{a_m}\max(|x_k^{(g)}|, |x_k^{(h)}|)$$

$$(k = m+1, \cdots, n).$$

Hence the index μ of this non-existing limit (8) is $\leqq m - 1$. For this index,

$$\lim_{\substack{g\to\infty\\h\to\infty}} x_\mu^{(g,h)}$$

either does not exist or exists and is different from zero. Hence there is an infinite one-dimensional sub-sequence

(9) $$X^{(g_i,h_i)} \qquad (i = 1, 2, 3, \cdots)$$

of the double sequence $X^{(g,h)}$, such that for all i

$$|x_\mu^{(g_i,h_i)}| \geqq c,$$

where c is a positive constant. Further obviously

$$\lim_{i\to\infty} F(X^{(g_i,h_i)} = 0,$$

$$\lim_{i\to\infty} |x_k^{(g_i,h_i)}| = 0 \qquad (k = m, m+1, \cdots, n),$$

and all $m - 1$ first coordinates

$$x_k^{(g_i,h_i)} \qquad (k = 1, 2, \cdots, m-1)$$

are bounded for $i \to \infty$.

Let ξ_i, for every i, be the coordinate

$$x_k^{(g_i,h_i)} \qquad (k = 1, 2, \cdots, m-1)$$

of maximum value $|x_k^{(g_i,h_i)}|$; hence

$$|\xi_i| \geqq c, \qquad \text{since} \qquad |\xi_i| \geqq |x_\mu^{(g_i,h_i)}|.$$

Then there is an infinite subsequence

$$X^{(g_{i_j},h_{i_j})} \qquad (j = 1, 2, 3, \cdots)$$

of the sequence (9), such that, if

$$X'^{(j)} = \frac{X^{(g_{i_j},h_{i_j})}}{\xi_{i_j}} = (x_1'^{(j)}, \cdots, x_n'^{(j)}) \qquad (j = 1, 2, 3, \cdots),$$

then all n limits

$$\lim_{j\to\infty} | x_k'^{(j)} | = a_k' \qquad (k = 1, 2, \cdots, n)$$

exist and satisfy the equations

$$\max (a_1', \cdots, a_n') = 1, \qquad a_m = a_{m+1} = \cdots = a_n = 0,$$

and

$$0 \leqq \lim_{j\to\infty} F(X'^{(j)}) \leqq \frac{1}{C} \lim_{j\to\infty} F(X^{(a_{i_j}, h_{i_j})}) = 0, \quad \text{i.e.} \quad \lim_{j\to\infty} F(X'^{(j)}) = 0,$$

Therefore the new sequence S'

$$X'^{(1)}, X'^{(2)}, X'^{(3)}, \cdots$$

has the same properties as S, but is of lower rank. Hence by induction with respect to the rank, a contradiction follows also in this case.—

By the inequality (6), all points of the convex body $C(\tau)$ lie in the finite cube

$$| X | \leqq \frac{\tau}{\gamma};$$

a convex body is therefore bounded. Conversely, if a convex set is bounded, then it is a convex body. For if its distance function $F(X)$ is not special, then there is at least one point $X^{(0)} \neq 0$, such that $F(X^{(0)}) = 0$; hence all points of the straight line passing through $X^{(0)}$ and the origin 0 belong to the set.

3. The character of a convex body. Let $C(\tau)$ be a convex body, $F(X)$ its distance function. If $X' \neq 0$ is an arbitrary vector, then the point $X = aX'$, where a is an element of $\mathfrak{K}$, lies in $C(\tau)$ provided that $| a |$ is either sufficiently small and positive, or 0. Hence for every index $h = 1, 2, \cdots, n$, the set S_h of all points

$$X = (x_1, \cdots, x_n) \qquad \text{with} \qquad x_1 = \cdots = x_{h-1} = 0, \qquad x_h \neq 0$$

of $C(\tau)$ is not empty and contains an infinity of elements. By (6),

$$| x_h | \leqq \frac{\tau}{\gamma}$$

for the points of S_h. Therefore $| x_h |$ has a positive upper bound ξ_h in this set, and to every $\epsilon > 0$ there is a point

$$X_\epsilon^{(h)} = (x_{1\epsilon}^{(h)}, \cdots, x_{n\epsilon}^{(h)}),$$

for which

$$F(X_\epsilon^{(h)}) \leqq \tau, \qquad x_{1\epsilon}^{(h)} = \cdots = x_{h-1\,\epsilon}^{(h)} = 0, \qquad \frac{\xi_h}{1+\epsilon} < | x_{h\epsilon}^{(h)} | \leqq \xi_h,$$

whereas there is no point X for which

$$F(X) \leqq \tau, \qquad x_1 = \cdots = x_{h-1} = 0, \qquad |x_h| > \xi_h .$$

The system of the n points

$$X_\epsilon^{(1)}, X_\epsilon^{(2)}, \cdots, X_\epsilon^{(n)}$$

corresponding to ϵ is obviously $\mathfrak{R}$-independent, and any point X of P_n can be written as

$$X = u_{1\epsilon}X_\epsilon^{(1)} + \cdots + u_{n\epsilon}X_\epsilon^{(n)},$$

where the u's belong to $\mathfrak{R}$ and are given explicitly by

$$u_{h\epsilon} = \sum_{k=1}^{n} \alpha_{hk\epsilon} x_k \qquad (h = 1, 2, \cdots, n)$$

with a matrix

$$(\alpha_{hk\epsilon})_{h,k=1,2,\cdots,n}$$

of non-vanishing determinant and elements depending on ϵ, but not on X.

We distinguish now whether the valuation $|x|$ of $\mathfrak{R}$, is *discrete* or not.

If $|x|$ is discrete, then there is a constant $b > 1$, such that for all $x \neq 0$ in $\mathfrak{R}$[4]

$$|x| = b^g$$

[4] If $|x|$ is discreet, then $F(X)$ has a similar property: *The set s of its values for X in P_n has no point of accumulation except* 0. This is clear for $n = 1$, for then all vectors are multipla of the unit vector (1). Suppose that the statement has already been proved for all spaces of $n - 1$ dimensions, but that it is not true in P_n. There is therefore an infinite sequence Σ of points

$$X^{(k)} = (x_1^{(k)}, \cdots, x_n^{(k)}) \qquad (k = 1, 2, 3, \cdots)$$

in P_n, such that all numbers

$$F(X^{(1)}), \qquad F(X^{(2)}), \qquad F(X^{(3)}), \cdots$$

are different from each other, and that the limit

$$\lim_{k\to\infty} F(X^{(k)}) = \lambda$$

exists and is positive. Write

$$X^{(k)} = x_1^{(k)} E^{(1)} + X^{(k)*} \qquad (k = 1, 2, 3, \cdots)$$

where

$$X^{(k)*} = (0, x_2^{(k)}, \cdots, x_n^{(k)}) \qquad (k = 1, 2, 3, \cdots)$$

lies in the $(n - 1)$-dimensional subspace P_{n-1}: $x_1 = 0$, of P_n. By (6), $|x_1^{(k)}|$ is bounded in Σ; hence we may assume that

$$\lim_{k\to\infty} |x_1^{(k)}| = \mu,$$

with a rational integer g depending on x. In this case the set of values $|x_{h\epsilon}^{(h)}|$ satisfies the equations

$$|x_{h\epsilon}^{(h)}| = \xi_h \qquad (h = 1, 2, \cdots, n)$$

for all sufficiently small ϵ. We assume that ϵ is sufficiently small and omit the index ϵ. Put

$$\Phi_\tau(X) = \tau \max(|u_1|, \cdots, |u_n|) = \tau \max_{h=1,2,\cdots,n} \left(\left| \sum_{k=1}^{n} \alpha_{hk} x_k \right| \right).$$

Then obviously

$$F(X) \leqq \tau, \quad \text{if} \quad \Phi_\tau(X) \leqq \tau.$$

Conversely let X be any point in $C(\tau)$. Then

$$|x_1| \leqq \xi_1$$

and therefore

$$|u_1| = \frac{|x_1|}{|x_1^{(1)}|} \leqq 1.$$

---

since, if necessary, we can replace Σ by an infinite subsequence. If $\mu = 0$, then for all sufficiently large k

$$F(X^{(k)}) = F(X^{(k*)}),$$

so that the sequence $X^{(1)*}, X^{(2)*}, X^{(3)*}, \cdots$ has the same properties as Σ, contrary to the hypothesis on P_{n-1}.

Hence if

$$\frac{x_1^{(k+1)}}{x_1^{(k)}} = q^{(k)}, \quad \text{then} \quad \lim_{k\to\infty} |q^{(k)}| = 1,$$

so that for all sufficiently large k

$$|q^{(k)}| = 1.$$

Obviously

$$X^{*(k)} = X^{(k+1)} - q^{(k)}X^{(k)} = X^{(k+1)*} - q^{(k)}X^{(k)*}$$

lies in P_{n-1}, and for all large k

$$F(X^{(k)}) = F(q^{(k)}X^{(k)}) \neq F(X^{(k+1)}).$$

Hence

$$F(X^{*(k)}) = \max(F(X^{(k)}), F(X^{(k+1)})).$$

Therefore the sequence of positive numbers

$$F(X^{*(1)}), F(X^{*(2)}), F(X^{*(2)}), \cdots$$

contains an infinity of different elements and has the limit λ, so that again a contradiction is obtained.

Hence, if

$$X' = X - u_1 X^{(1)} = (0, x_2', \cdots, x_n'),$$

then

$$F(X') \leqq \max (F(X), | u_1 | F(X^{(1)})) \leqq \tau,$$

and so X' also belongs to $C(\tau)$. Therefore

$$| x_2' | \leqq \xi_2 ,$$

so that

$$| u_2 | = \frac{| x_2' |}{| x_2^{(2)} |} \leqq 1.$$

Continuing in this way, we obtain all inequalities

$$| u_1 | \leqq 1, \cdots , | u_n | \leqq 1,$$

i.e. we have proved

$$\Phi_\tau(X) \leqq t, \qquad \text{if} \qquad F(X) \leqq \tau.$$

The domain defined by

$$\frac{1}{\tau} \Phi_\tau(X) = \max_{h=1,2,\cdots,n} \left(\left| \sum_{k=1}^{n} \alpha_{hk} x_k \right| \right) \leqq 1$$

is called a parallelepiped; our result may therefore be expressed in the form: *If the valuation $| x |$ is discrete, than every convex body $C(\tau)$ is a parallelepiped.*

As we have proved, the two domains

$$F(X) \leqq \tau \qquad \text{and} \qquad \Phi_\tau(X) \leqq \tau$$

are identical. In general, this does not imply the identity[5]

$$F(X) = \Phi_\tau(X)$$

for all X, and the function $\Phi_\tau(X)$ depends on τ. Suppose, however, that the set of values of $F(X)$ is the same as that of the values of $| x |$, and that τ is also an element of this set.[6] Then

$$\Phi_\tau(X) = \Phi(X)$$

becomes independent of τ, and for all X in P_n identically

$$F(X) = \Phi(X), \tag{10}$$

as follows easily from the property (B) of the distance functions.—

---

[5] E.g., if $\mathfrak{R} = \mathfrak{K}$ is the p-adic field ($p \geqq 3$), $n = 2$, and

$$F(X) = \max (| x_1 |_p , 2 | x_2 |_p) .$$

[6] It suffices to assume that $F(X)$ does not assume every positive value, and that the equation $F(X) = \tau$ has no solution.

Next assume that *the valuation* $|x|$ *is not discrete*, so that *its values lie everywhere dense on the positive real axis.* Now the n vectors

$$X_\epsilon^{(1)}, X_\epsilon^{(2)}, \cdots, X_\epsilon^{(n)}$$

will depend on ϵ, and so does the function

$$\Phi_{\tau\epsilon}(X) = \tau \max_{h=1,2,\cdots,n} (|u_{h\epsilon}|) = \tau \max_{h=1,2,\cdots,n} \left(\left|\sum_{k=1}^{n} \alpha_{hk\epsilon} x_k\right|\right).$$

Evidently

(11) $$F(X) \leqq \tau, \quad \text{if} \quad \Phi_{\tau\epsilon}(X) \leqq \tau.$$

Conversely, suppose $F(X) \leqq \tau$. Then

$$|x_1| \leqq \xi_1$$

and therefore

$$|u_{1\epsilon}| = \frac{|x_1|}{|x_{1\epsilon}^{(1)}|} < 1 + \epsilon.$$

Hence, if

$$X'_\epsilon = X - u_{1\epsilon} X_\epsilon^{(1)} = (0, x'_{2\epsilon}, \cdots, x'_{u\epsilon}),$$

then

$$F(X'_\epsilon) \leqq \max (F(X), |u_{1\epsilon}| F(X_\epsilon^{(1)})) < (1 + \epsilon)\tau.$$

There is a number α_ϵ in $\mathfrak{K}$ such that

$$F(X'_\epsilon) \leqq |\alpha_\epsilon| \tau \leqq (1 + \epsilon)\tau, \quad \text{i.e.} \quad F(\alpha_\epsilon^{-1} X'_\epsilon) \leqq \tau.$$

Hence

$$|\alpha_\epsilon^{-1} x'_{2\epsilon}| \leqq \xi_2, \qquad |x'_{2\epsilon}| \leqq (1 + \epsilon)\xi_2,$$

and therefore

$$|u_{2\epsilon}| = \frac{|x'_{2\epsilon}|}{|x_{2\epsilon}^{(2)}|} < (1 + \epsilon)^2,$$

so that, if

$$X''_\epsilon = X'_\epsilon - u_{2\epsilon} X_\epsilon^{(2)} = X - (u_{1\epsilon} X_\epsilon^{(1)} + u_{2\epsilon} X_\epsilon^{(2)}) = (0, 0, x''_{3\epsilon}, \cdots, x''_{n\epsilon}),$$

then

$$F(X''_\epsilon) \leqq \max (F(X'_\epsilon), |u_{2\epsilon}| F(X_\epsilon^{(2)})) < (1 + \epsilon)^2 \tau.$$

Continuing in the same way, we obtain the n inequalities

$$|u_{h\epsilon}| < (1 + \epsilon)^h \qquad (h = 1, 2, \cdots, n),$$

hence

(12) $$\Phi_{\tau\epsilon}(X) < (1 + \epsilon)^n \tau, \quad \text{if} \quad F(X) \leqq \tau.$$

From (11) and (12), since $\epsilon > 0$ is arbitrarily small:

If the valuation $|x|$ is everywhere dense on the positive axis, then the convex body $C(\tau)$ can be approximated arbitrarily near both from the inside and outside by means of parallelepipeds.

Take now, say $\tau = 1$ and put

$$\Phi_\epsilon(X) = \Phi_{1\epsilon}(X) = \max_{h=1,2,\cdots,n}\left(\left|\sum_{k=1}^{n} \alpha_{hk\epsilon} x_k\right|\right).$$

To every point X, there are two elements α and β of $\mathfrak{K}$, such that

$$\Phi_\epsilon(X) \leqq |\alpha| \leqq (1+\epsilon)\Phi_\epsilon(X) \quad \text{and} \quad F(X) \leqq |\beta| \leqq (1+\epsilon)F(X).$$

Hence from (11)

$$\Phi_\epsilon\left(\frac{X}{\alpha}\right) \leqq 1, \qquad F\left(\frac{X}{\alpha}\right) \leqq 1, \qquad F(X) \leqq |\alpha| \leqq (1+\epsilon)\Phi_\epsilon(X),$$

and from (12)

$$F\left(\frac{X}{\beta}\right) \leqq 1, \quad \Phi_\epsilon\left(\frac{X}{\beta}\right) \leqq (1+\epsilon)^n, \quad \Phi_\epsilon(X) \leqq (1+\epsilon)^n|\beta| \leqq (1+\epsilon)^{n+1}F(X),$$

and therefore uniformly in X

$$(1+\epsilon)^{-(n+1)}\Phi_\epsilon(X) \leqq F(X) \leqq (1+\epsilon)\Phi_\epsilon(X). \tag{13}$$

In general, these inequalities cannot be improved to an equation analogous to (10), e.g. if $F(X) = \tau$ has no solution.

4. The character of a convex set. If $F(X)$ is not special, then the set M of all solutions of $F(X) = 0$ contains elements other than $X = 0$. From (B) and (C), with X and Y also $aX + bY$ belongs to M, if a and b are elements of $\mathfrak{K}$. Hence M is a $\mathfrak{K}$-modul, say of dimension $n - m$. Obviously $m < n$; it is possible that $m = 0$, but then $F(X)$ vanishes identically and $C(\tau)$ is the whole space. Suppose therefore, that $1 \leqq m \leqq n - 1$, and let

$$P^{(m+1)}, P^{(m+2)}, \cdots, P^{(n)}$$

be $n - m\mathfrak{K}$-independent elements of M,

$$P^{(1)}, P^{(2)}, \cdots, P^{(m)}$$

m other points of P_n, so that the system of n vectors

$$P^{(1)}, P^{(2)}, \cdots, P^{(n)}$$

is still $\mathfrak{K}$-independent. Then every point X in P_n can be written as

$$X = v_1P^{(1)} + \cdots + v_nP^{(n)}$$

with elements $v_1, \cdots, v_n$ of $\mathfrak{K}$, viz.

$$v_h = \sum_{k=1}^{n} \beta_{hk}x_k \qquad (h = 1, 2, \cdots, n),$$

where the constant matrix in $\mathfrak{R}$

$$(\beta_{hk})_{h,k=1,2,\dots,n}$$

has non-vanishing determinant. Since

$$F\left(\sum_{h=m+1}^{n} v_k P^{(h)}\right) = 0,$$

we have

$$F(X) = F\left(\sum_{h=1}^{m} v_h P^{(h)}\right) = \Psi(V),$$

where

$$\Psi(V) = \Psi(v_1, \cdots, v_m) = \Psi\left(\sum_{k=1}^{n} \beta_{1k} x_k, \cdots, \sum_{k=1}^{n} \beta_{nk} x_k\right)$$

is now obviously a special distance function in the m-dimensional space P_m of all points $V = (v_1, \cdots, v_m)$. Every convex set with $m > 0$ can therefore be considered as a cylinder, the basis of which is a convex body of $m < n$ dimensions.

5. The polar body of $C(\tau)$. Let $F(X)$ be the general distance function of §4, Y an arbitrary vector in P_n. Then we define a function $G(Y)$ by

(14) $G(0) = 0$; $G(Y) = \limsup (|XY|)$ *for all* X *with* $F(X) \leqq 1$, *if* $Y \neq 0$.

In order to determine this function, let

$$Q^{(1)}, Q^{(2)}, \cdots, Q^{(n)}$$

be the n points in P_n, which satisfy the equations

$$P^{(h)} Q^{(k)} = \begin{cases} 1 & \text{for } h = k, \\ 0 & \text{for } h \neq k, \end{cases}$$

and write

$$Y = w_1 Q^{(1)} + \cdots + w_n Q^{(n)};$$

then

$$w_h = \sum_{k=1}^{n} \gamma_{hk} y_k \qquad (h = 1, 2, \cdots, n),$$

where the determinant of the matrix in $\mathfrak{R}$

$$(\gamma_{hk})_{h,k=1,2,\dots,n}$$

does not vanish. Then

$$XY = v_1 w_1 + \cdots + v_n w_n.$$

Hence obviously

$$G(Y) = \infty, \qquad \text{unless} \qquad w_{m+1} = \cdots = w_n = 0.$$

Suppose therefore that

$$(15) \qquad w_{m+1} = w_{m+2} = \cdots = w_n = 0,$$

and put

$$G(Y) = \mathrm{X}(W),$$

where $W = (w_1, \cdots, w_m)$ is a vector in P_m. Then from (14),

$$(16) \quad \mathrm{X}(0) = 0; \quad \mathrm{X}(W) = \limsup\,(|\,VW\,|) \textit{ for all } V \textit{ with } \Psi(V) \leqq 1, \quad \textit{if } W \neq 0,$$

so that the relation of $\mathrm{X}(W)$ to $\Psi(V)$ is the same as that of $G(Y)$ to $F(X)$. By §4, $\Psi(V)$ is a *special* distance function, and so is $\mathrm{X}(W)$, as follows easily from (16) and the properties (A′), (B), and (C) of $\Psi(V)$.

We call $G(Y)$ the *polar function to* $F(X)$; *for* $m < n$ *it is not itself a distance function*, but becomes one in the m-dimensional space (15), where it coincides with $\mathrm{X}(W)$. The set $C'(1/\tau)$: $G(Y) \leqq 1/\tau$, is further called the *polar set to* $C(\tau)$; it lies entirely in (15) and here is identical with the convex body $\mathrm{X}(W) = 1/\tau$.

Suppose now that $m = n$, i.e. both $F(X)$ and $G(Y)$ are special distance functions; then the polar set $C'(1/\tau)$ becomes a convex body. We shall prove that in this case *the relation between* $F(X)$ *and* $G(Y)$ *is reciprocal*, i.e. $F(X)$ is the polar function to $G(Y)$ and $C(\tau)$ the polar body to $C'(1/\tau)$.

This assertion is evident, if $F(X) = |\,X\,|$, for then obviously $G(Y) = |Y|$. Further let

$$\Omega = (a_{hk})_{h,k=1,2,\ldots,n}\,; \qquad \Omega^K = (a_{hk}^K)_{h,k=1,2,\ldots,n}$$

be an arbitrary matrix in $\mathfrak{K}$ with nonvanishing determinant, and its complementary matrix, so that for all X and Y the scalar product[7]

$$\Omega X \cdot \Omega^K Y = XY.$$

Then the transformed distance functions $G'(Y) = G(\Omega^K Y)$ and $F'(X) = F(\Omega X)$ have still the property that the first one is polar to the second, since

$$G'(Y) = G(\Omega^K Y) = \limsup_{F(X) \leqq 1}\,(|\,X \cdot \Omega^K Y\,|)$$

$$= \limsup_{F(\Omega X) \leqq 1}\,(|\,\Omega X \cdot \Omega^K Y\,|) = \limsup_{F'(X) \leqq 1}\,(|\,XY\,|).$$

Further, if $F_1(X)$ and $F_2(X)$ are two distance functions such that for all X

$$F_1(X) \leqq F_2(X),$$

---

[7] The vector $X' = (x_1', \cdots, x_n') = \Omega X$ is defined by $x_h' = \sum_{k=1}^n a_{hk}x_k$ for $h = 1, 2, \cdots, n$.

then the polar distance functions $G_1(Y)$ and $G_2(Y)$ satisfy the inverted inequality

$$G_1(Y) \geqq G_2(Y).$$

We distinguish now the same two cases as in §3. If the valuation $|x|$ is discrete, then we showed the existence of a matrix

$$A = (\alpha_{hk})_{h,k=1,2,\ldots,n}$$

in $\mathfrak{K}$ with determinant different from zero, such that

$$F(X) = \Phi(X) = |AX|$$

identically in X. The polar function to $F(X)$ is therefore

$$G(Y) = |A^K Y|,$$

and since $(A^K)^K = A$, the statement follows at once.—In this case, the definition of $G(Y)$ can obviously be replaced by the simpler one:

$$G(Y) = \max_{X \neq 0} \frac{|XY|}{F(X)}. \tag{17}$$

Secondly, let $|x|$ be everywhere dense on the positive real axis. Then to every $\delta > 0$, there are two matrices

$$A_1 = (\alpha_{hk}^{(1)})_{h,k=1,2,\ldots,n} \quad \text{and} \quad A_2 = (\alpha_{hk}^{(2)})_{h,k=1,2,\ldots,n}$$

in $\mathfrak{K}$ with non-vanishing determinants, such that if

$$F_1(X) = |A_1X|, \qquad F_2(X) = |A_2X|,$$

then for all X

$$F_1(X) \leqq F(X) \leqq F_2(X) \leqq (1 + \delta)F_1(X),$$

as follows easily from (13). Hence if

$$G_1(Y) = |A_1^K Y|, \qquad G_2(Y) = |A_2^K Y|$$

are the polar functions to $F(X)$ and $F(X)$, then also

$$G_2(Y) \leqq G(Y) \leqq G_1(Y);$$

and[8]

$$G_2(Y) \leqq (1 + 2\delta)G_1(Y),$$

---

[8] There is a number α in $\mathfrak{K}$ such that

$$1 + \delta \leqq |\alpha| \leqq 1 + 2\delta.$$

Then by hypothesis

$$F_1(X) \leqq (1 + \delta)F_2(X) \leqq F_2(\alpha X).$$

Hence

$$\frac{1}{1 + 2\delta} G_2(Y) \leqq G_2\left(\frac{Y}{\alpha}\right) \leqq G_1(Y),$$

since the polar function to $F_2(\alpha X)$ is $G_2\left(\frac{Y}{\alpha}\right)$.

for all Y. Since δ can be taken arbitrarily small, the assertion follows again for the same reason.—In this case, the definition of $G(Y)$ is easily replaced by

$$G(Y) = \lim_{X \neq 0} \sup \frac{|XY|}{F(X)}. \tag{17'}$$

By the proved reciprocity of $F(X)$ and $G(Y)$, the formulae (17) and (17′) remain true if $G(Y)$ is replaced by $F(X)$ and vice versa.

II. "Geometry of Numbers" in a Domain of Power Series

6. Notation. We specialize now the fields $\mathfrak{R}$ and $\mathfrak{K}$ of §1, and denote by

$\mathfrak{k}$ *an arbitrary field,*

z *an indeterminate,*

$\mathfrak{T} = k[z]$ *the ring of all polynomials in z with coefficients in* $\mathfrak{k}$,

$\mathfrak{R} = k(z)$ *the quotient field of* $\mathfrak{T}$, i.e. *the field of all rational functions in z with coefficients in* $\mathfrak{k}$,

$|x|$ *the special valuation of* $\mathfrak{R}$ *defined by*

$$|x| = \begin{cases} 0, & \text{if } x = 0, \\ e^f, & \text{if } x \neq 0 \text{ is of order } f,^9 \end{cases}$$

$\mathfrak{K}$ *the perfect extension of* $\mathfrak{R}$ *with respect to this valuation, i.e. the field of all formal Laurent series*

$$x = \alpha_f z^f + \alpha_{f-1} z^{f-1} + \alpha_{f-2} z^{f-2} + \cdots$$

with coefficients in $\mathfrak{k}$; *if* α_f *is the non-vanishing coefficient with highest index* $\gtreqless 0$, *then* $|x| = e^f$,

Λ_n *the set of all "lattice points" in* P_n, *i.e. that of all points with coordinates in* $\mathfrak{T}$.

The valuation $|x|$ is by definition a power of e with integral exponent. We assume the same for all distance functions which we consider from now onwards, and we shall consider only convex sets or bodies $C(\tau)$, where τ is an exact power of e, say $\tau = e^t$.

7. The volume V of a convex body $C(1)$. Let $F(X)$ be a special distance function, $C(e^t)$ the convex body $F(X) \leqq e^t$, where t is an arbitrary integer. It is obvious that the set $\mathfrak{m}(t)$ of all lattice points in $C(e^t)$ forms a $\mathfrak{k}$-modul. In the special case $F(X) = |X|$, this set has exactly

$$M_0(t) = n(t + 1)$$

$\mathfrak{k}$-independent elements. Hence, by the inequalities (5) and (6), $\mathfrak{m}(t)$ has always a finite dimension $M(t)$, and this dimension is certainly positive for large t.

---

[9] The order of a rational function is the degree of its numerator minus the degree of its denominator.

Obviously

$$M_0(t + 1) = M_0(t) + n. \tag{18}$$

Suppose that t is already so large that

$$e^{t+1} \geqq \Gamma.$$

Then a lattice point in $C(e^{t+1})$ can be written as

$$X = X_0 + zX_1,$$

where X_0 and X_1 are again lattice points, and the coordinates of X_0 lie in $\mathfrak{k}$, i.e.

$$| X_0 | \leqq 1, \qquad F(X_0) \leqq \Gamma \leqq e^{t+1}.$$

Hence

$$F(zX_1) \leqq \max\,(F(X), F(X_0)) \leqq e^{t+1}, \qquad F(X_1) \leqq e^t,$$

so that X_1 lies in $\mathfrak{m}(t)$. Conversely, if X_1 belongs to $\mathfrak{m}(t)$, then

$$F(X) \leqq \max\,(F(zX_1), F(X_0)) \leqq e^{t+1}.$$

Now the two vectors X_0 and zX_1, where X_0 and X_1 are lattice points and $| x_0 | \leqq 1$, are $\mathfrak{k}$-independent, and the X_0 form a $\mathfrak{k}$-modul of dimension n. Hence

$$M(t + 1) = M(t) + n. \tag{19}$$

The two equations (18) and (19) show that for large t, the function $M(t) - M_0(t)$ of t is independent of t. Hence the limit

$$V = \lim_{t\to\infty} e^{M(t) - M_0(t)} \tag{20}$$

exists; it is called the *volume of the convex body* $C(1)$.[10] In particular, if $F(X) = |X|$, then obviously $V = 1$.

8. The invariance of V. Let

$$\Omega = (a_{hk})_{h,k=1,2,\ldots,n} \quad \text{and} \quad \Omega^I = (a^I_{hk})_{h,k=1,2,\ldots,n}$$

be a matrix with elements in $\mathfrak{K}$ and determinant $D \neq 0$, and its inverse matrix. The linear transformation

$$Y = \Omega X \quad \text{or} \quad X = \Omega^I Y$$

changes $F(X)$ into the new distance function

$$F'(Y) = F(X) = F(\Omega^I Y);$$

let $C'(e^t)$ be the corresponding convex body $F'(Y) \leqq e^t$, and V' the volume of $C'(1)$. *Then*

$$V' = | D | V. \tag{21}$$

[10] This definition is analogous to that of the volume of a body by means of lattice points in an ordinary real space.

PROOF: We denote by $\mathfrak{m}'(t)$ the $\mathfrak{k}$-modul of all lattice points in $C'(e^t)$, by $M'(t)$ the dimension of $\mathfrak{m}'(t)$, and prove the statement in a number of steps.

1: The elements of Ω lie in $\mathfrak{T}$, and D belongs to $\mathfrak{k}$.

The formulae $Y = \Omega X$, $X = \Omega^I Y$ establish a (1, 1)-correspondence between the elements X of $\mathfrak{m}(t)$ and Y of $\mathfrak{m}'(t)$. Obviously, this correspondence changes every linear relation

$$\alpha_1 X^{(1)} + \cdots + \alpha_r X^{(r)} = 0$$

with coefficients in $\mathfrak{k}$ into the identical relation in the Y's, and vice versa; therefore $\mathfrak{k}$-independent elements of $\mathfrak{m}(t)$ or $\mathfrak{m}'(t)$ are transformed into $\mathfrak{k}$-independent members of the other modul. Hence both moduls have the same dimension: $M(t) = M'(t)$, q.e.d.

2: Ω is a triangle matrix

$$\Omega = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & 0 \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

with elements in $\mathfrak{T}$ and determinant

$$D = a_{11}a_{22} \cdots a_{nn} \neq 0.$$

The equation $Y = \Omega X$ denotes that

$$\begin{aligned} y_1 &= a_{11}x_1, \\ y_2 &= a_{21}x_1 + a_{22}x_2, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n; \end{aligned}$$

hence every lattice point Y can be written as[11]

$$Y = \Omega X^* + Y^*,$$

where X^* and Y^* are again lattice points and $Y^* = (y_1^*, \cdots, y_n^*)$ satisfies the inequalities

$$|y_1^*| < |a_{11}|, \; |y_2^*| < |a_{22}|, \cdots, |y_n^*| < |a_{nn}|.$$

Therefore

$$|Y^*| \leqq c_1, \text{ i.e. } F'(Y^*) \leqq c_1\Gamma',$$

where c_1 is a positive constant depending only on Ω, and Γ' is the constant in (5) belonging to $F'(Y)$. The set of all vectors Y^* forms a $\mathfrak{k}$-modul $\mathfrak{m}^*$ of dimension d, where

$$e^d = |a_{11}||a_{22}| \cdots |a_{nn}| = |D|.$$

---

[11] We use the trivial lemma: "*To a and $b = 0$ in $\mathfrak{T}$ there is a q and an r in $\mathfrak{T}$, such that $a = bq + r$ and $|r| < |b|$.*

Let t be so large that

$$e^t \geqq c_1\Gamma'.$$

Then for X^* in $\mathfrak{m}(t)$

$$F'(Y) = F(\Omega^I Y) = F(X^* + \Omega^I Y^*) \leqq \max\,(F(X^*), F'(Y^*)) \leqq e^t,$$

and conversely for Y in $\mathfrak{m}'(t)$

$$F(X^* + \Omega^I Y^*) \leqq e^t, \text{ i.e. } F(X^*) \leqq \max\,(F(X^* + \Omega^I Y^*), F'(Y^*)) \leqq e^t.$$

There is therefore a (1, 1)-correspondence between the elements Y of $\mathfrak{m}'(t)$ and the pairs (X^*, Y^*) of one element X^* of $\mathfrak{m}(t)$ and one element Y^* of $\mathfrak{m}^*$. Hence $M'(t) = M(t) + d$, q.e.d.

3: The elements of Ω belong to $\mathfrak{T}$.

The result follows immediately from the two previous steps, since Ω, as is well known,[12] can be written as $\Omega = \Omega_1\Omega_2$, where the two factors are of the classes 1 and 2.

4: The elements of Ω lie in $\mathfrak{N}$.

Now $\Omega = \Omega_a\Omega_b^I$, where both Ω_a and Ω_b are of the class 3, so that the statement follows at once.

5: Ω has elements in $\mathfrak{K}$, such that

$$|\,D\,| = 1, \qquad |\,a_{hk}\,| \leqq 1 \qquad (h, k = 1, 2, \cdots, n).$$

Then the same inequalities hold for the inverse matrix Ω^I, so that for every point X

$$|\,\Omega X\,| \leqq |\,X\,|, \qquad |\,X\,| = |\,\Omega^I\Omega X\,| \leqq |\,\Omega X\,|,$$

and therefore

$$|\,X\,| = |\,\Omega X\,| = |\,\Omega^I X\,|.$$

Now to every lattice point X there is a second lattice point Y such that with a suitable point Y^*

$$\Omega X = Y + Y^*, \qquad |\,Y^*\,| < 1;$$

then conversely

$$\Omega^I Y = X + X^*, \qquad |\,X^*\,| < 1,$$

and

$$X^* = -\Omega^I Y^*, \qquad \Omega X^* = -Y^*.$$

The relation between X and Y is therefore a (1, 1)-correspondence which obviously leaves invariant the property of $\mathfrak{k}$-independence. Suppose that

$$e^t \geqq \Gamma.$$

[12] This can be proved, e.g. by a method analogous to Minkowski's "adaptation" of a lattice; Geometrie der Zahlen §46.

Then for X in $\mathfrak{m}(t)$

$$F(X^*) < \Gamma \leqq e^t,$$

and therefore

$$F'(Y) = F(\Omega^I Y) = F(X + X^*) \leqq \max (F(X), F(X^*)) \leqq e^t,$$

so that Y lies in $\mathfrak{m}'(t)$; conversely, if Y belongs to $\mathfrak{m}'(t)$, then X is an element of $\mathfrak{m}(t)$. Hence $M(t) = M'(t)$, q.e.d.

6: Finally, let Ω have elements in $\mathfrak{K}$. Then it can be split into

$$\Omega = \Omega_4 + \Omega^*$$

where Ω_4 is of the class 4, while the elements of Ω^* lie in $\mathfrak{K}$ and have so small values that

$$\Omega_5 = \Omega_4^I \Omega$$

is of the class 5. Then the result follows at once, since $\Omega = \Omega_4 \Omega_5$.

Two conclusions are immediate from (21). The convex body $C(e^t)$, i.e. $F(z^{-t}X') \leqq 1$, is obtained from $C(1)$ by the transformation $X' = z^t X$; hence it has the volume $V(e^t) = e^{nt}V$. Secondly, *let $G(Y)$ be the polar distance function to $F(X)$, and V' the volume of the convex body $C'(1)$*, i.e. $G(Y) \leqq 1$. *Then V and V' are related by the equation*

$$VV' = 1. \tag{22}$$

For by §5, there is a matrix A with non-vanishing determinant, such that

$$F(X) = | \mathrm{A}X | \quad \text{and} \quad G(Y) = | \mathrm{A}^K Y |,$$

hence

$$V = (| \mathrm{A} |)^{-1} \quad \text{and} \quad V' = (| \mathrm{A}^k |)^{-1} = | \mathrm{A} |;$$

the statement is therefore obvious.

9. The minima of $F(X)$. To the distance function $F(X)$, there exist n $\mathfrak{K}$-independent lattice points

$$X^{(k)} = (x_1^{(k)}, \cdots, x_n^{(k)}) \qquad (k = 1, 2, \cdots, n),$$

such that

$F(X^{(1)}) = \sigma^{(1)} = e^{g_1}$ *is the minimum of $F(X)$ in all lattice points $X \neq 0$,*

$F(X^{(2)}) = \sigma^{(2)} = e^{g_2}$ *is the minimum of $F(X)$ in all lattice points X which are $\mathfrak{K}$-independent of $X^{(1)}$*, etc., and finally

$F(X^{(n)}) = \sigma^{(n)} = e^{g_n}$ *is the minimum of $F(X)$ in all lattice points X which are $\mathfrak{K}$-independent of $X^{(1)}, X^{(2)}, \cdots, X^{(n-1)}$.*

The numbers $\sigma^{(1)}, \sigma^{(2)}, \cdots, \sigma^{(n)}$ are called *the n successive minima of $F(X)$*. By this construction, the determinant

$$D = | x_h^{(k)} |_{h,k=1,2,\cdots,n}$$

lies in $\mathfrak{T}$ and does not vanish; further obviously

(23) $$0 < \sigma^{(1)} \leqq \sigma^{(2)} \leqq \cdots \leqq \sigma^{(n)} \quad \text{and} \quad g_1 \leqq g_2 \leqq \cdots \leqq g_n .$$

We shall prove the two equations

(24) $$| D | = 1,$$

(25) $$\sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)} = \frac{1}{V};$$

in the second one, V is again the volume of $C(1)$. Thus, in particular, *D is an element of $\mathfrak{k}$, and may obviously be taken as equal to* 1.

A: Proof of (24). Every point X in P_n can be written as

$$X = y_1X^{(1)} + \cdots + y_nX^{(n)},$$

where the y's are elements of $\mathfrak{K}$. Then the coordinates x_h of X are linear functions with determinant D of the coordinates y_h of $Y = (y_1, \cdots, y_n)$. We define a new distance function $\Pi(X)$ by

$$\Pi(X) = | Y |.$$

By (21), the convex body $\Pi(X) \leqq 1$ has the volume $| D |$; we determine it in the following way:

If X is a lattice point, then Y also has its coordinates y_h in $\mathfrak{T}$. For since with Y also X is obviously a lattice point, we may assume without loss of generality that

(26) $$\Pi(X) = | Y | < 1,$$

and have to show that no lattice point $X \neq 0$ satisfies this inequality. Let m, where $1 \leqq m \leqq n$, be the greatest index for which $y_m \neq 0$. Then

$$X = \sum_{h=1}^{m} y_hX^{(h)}, \qquad X^{(1)}, \cdots, X^{(m-1)}$$

are $\mathfrak{R}$-independent lattice points, and by (26)

$$F(X) \leqq \max (| y_1 |F(X^{(1)}), \cdots, | y_m | F(X^{(m)})) < \sigma^{(m)},$$

in contradiction to the minimum property of $\sigma^{(m)}$.

Hence there are exactly $M_1(t) = n(t + 1)$ $\mathfrak{k}$-independent lattice points such that $\Pi(X) \leqq e^t$, viz. all points corresponding to a basis of $\mathfrak{k}$-independent points Y with $| Y | \leqq e^t$. Therefore

$$| D | = \lim_{t\to\infty} e^{M_1(t)-M_0(t)} = 1, \qquad \text{q.e.d.}$$

B: Proof of (25). Now we use the fact that every point X in P_n can be written as

$$X = y_1z^{-g_1}X^{(1)} + \cdots + y_nz^{-g_n}X^{(n)},$$

where the y's belong to $\mathfrak{K}$. Let $\Sigma(X)$ be the distance function given by

$$\Sigma(X) = | Y |.$$

Since

$$F(z^{-g_h}X^{(h)}) = 1 \qquad (h = 1, 2, \cdots, n),$$

obviously

$$F(X) \leqq 1, \quad \text{if} \quad \Sigma(X) \leqq 1.$$

But the converse is also true: If

$$F(X) \leqq 1, \quad \text{then} \quad \Sigma(X) \leqq 1,$$

and therefore evidently

$$F(X) = \Sigma(X) = | Y |,$$

identically in X.

For suppose that on the contrary for a certain point X in P_n

$$F(X) \leqq 1, \quad \text{but} \quad \Sigma(X) > 1.$$

Then let m with $1 \leqq m \leqq n$ be the greatest index for which $| y_m | > 1$; hence if $m < n$

$$| y_{m+1} | \leqq 1, \cdots, | y_n | \leqq 1.$$

Write

$$y_h = zy_h^* + y_h^{**} \qquad (h = 1, 2, \cdots, n),$$

where the y_h^* are elements of $\mathfrak{T}$, the y_h^{**} elements of $\mathfrak{K}$, and

$$y_m^* \neq 0, \qquad y_{m+1}^* = \cdots = y_n^* = 0, \qquad | y_1^{**} | \leqq 1, \cdots, | y_n^{**} | \leqq 1,$$

and put

$$Y^* = (y_1^*, \cdots, y_n^*), \qquad Y^{**} = (y_1^{**}, \cdots, y_n^{**}),$$

so that

$$Y = zY^* + Y^{**}.$$

Obviously, Y^* is a lattice point, Y^{**} a point such that $| Y^{**} | \leqq 1$. Also write

$$X^* = \sum_{h=1}^{n} y_h^* z^{-g_h} X^{(h)} = \sum_{h=1}^{m} y_h^* z^{-g_h} X^{(h)}, \qquad X^{**} = \sum_{h=1}^{n} y_h^{**} z^{-g_h} X^{(h)},$$

so that

$$X = zX^* + X^{**}.$$

Then from $\Sigma(X^{**}) = | Y^{**} | \leqq 1$,

$$F(X^{**}) \leqq 1.$$

Hence

$$F(zX^*) \leqq \max (F(X), F(X^{**})) \leqq 1, \qquad F(X^*) < 1,$$

and

$$F(X^0) < \sigma^{(m)},$$

where $X^0 = z^{g_m}X^*$. This inequality, however, is impossible, since the m lattice points

$$X^0 = \sum_{h=1}^{m} y_h^* z^{g_m - g_h} X^{(h)}, \qquad X^{(1)}, \cdots, X^{(m-1)}$$

are $\mathfrak{K}$-independent, so that by the minimum property of $\sigma^{(m)}$

$$F(X^{(0)}) \geqq \sigma^{(m)}.$$

Therefore (27) is true, so that by the invariance theorem of §8

$$V = \frac{|D|}{\sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)}} = (\sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)})^{-1},$$

since the transformation of X into Y has the determinant

$$Dz^{-(g_1+g_2+\cdots+g_n)}.$$

The equation (25) is therefore proved.

From this equation and from (23) in particular

$$\sigma^{(1)} \leqq V^{-1/n};$$

i.e. *to every distance function* $F(X)$ *there is a lattice point* $X \neq 0$ *such that*

$$F(X) \leqq \frac{1}{\sqrt[n]{V}}.$$

Here equality holds if and only if all minima

$$\sigma^{(1)} = \sigma^{(2)} = \cdots = \sigma^{(n)},$$

thus certainly not, if V is not an integral power of e^n.

10. The relations between the minima of $F(X)$ and $G(Y)$. To the n lattice points $X^{(1)}, X^{(2)}, \cdots, X^{(n)}$ defined in the last paragraph, we construct n points $Y^{(1)}, Y^{(2)}, \cdots, Y^{(n)}$ satisfying

$$X^{(h)}Y^{(n-k+1)} = \begin{cases} 1 & \text{for } h = k, \\ 0 & \text{for } h \neq k; \end{cases} \tag{27}$$

since $|D| = 1$, these points are lattice points. We further define n positive numbers

$$\tau^{(h)} = \frac{1}{\sigma^{(n-h+1)}} = e^{j_h} \qquad (h = 1, 2, \cdots, n), \tag{28}$$

so that

$$(29) \qquad 0 < \tau^{(1)} \leqq \tau^{(2)} \leqq \cdots \leqq \tau^{(n)} \quad \text{and} \quad j_1 \leqq j_2 \leqq \cdots \leqq j_n .$$

Then $F(X)$ and the polar function $G(Y)$ can be written as

$$(30) \qquad F(X) = \max_{h=1,2,\cdots,n} (\sigma^{(h)} \mid XY^{(n-h+1)} \mid),$$

$$(31) \qquad G(Y) = \max_{h=1,2,\cdots,n} (\tau^{(h)} \mid YX^{(n-h+1)} \mid),$$

thus in an entirely symmetrical way. For we proved in the preceding paragraph that if X is written as

$$(32) \qquad X = \sum_{h=1}^{n} y_h z^{-g_h} X^{(h)},$$

then

$$F(X) = \mid Y \mid, \qquad Y = (y_1, y_2, \cdots, y_n).$$

But by multiplying (32) scalar with $Y^{(n)}, \cdots, Y^{(1)}$, we get by (27)

$$y_h = z^{g_h} \cdot (XY^{(n-h+1)}) \qquad (h = 1, 2, \cdots, n)$$

and therefore (30). The formula (31) is a consequence of (30) by the results in §5.[13]

From (27) and (31)

$$(33) \qquad G(Y^{(h)}) = \tau^{(h)} = e^{j_h}.$$

We prove now that these numbers $\tau^{(h)}$ in their natural order are the n successive minima of $G(Y)$ in Λ_n. Obviously it suffices to show that *if*

$$Z^{(1)}, Z^{(2)}, \cdots, Z^{(n)}.$$

are any n $\mathfrak{K}$-independent lattice points, such that

$$G(Z^{(1)}) \leqq G(Z^{(2)}) \leqq \cdots \leqq G(Z^{(n)}),$$

[13] We can prove (31) directly in the following way: Obviously

$$X = \sum_{h=1}^{n} (XY^{(n-h+1)}) X^{(h)},$$

where the brackets are again the scalar products. Hence from (14)

$$G(Y) = \max (\mid XY \mid) = \max \left(\left| \sum_{h=1}^{n} (XY^{(n-h+1)})(X^{(h)}Y) \right| \right),$$

where the maximum extends over all points X of $C(1)$, i.e. for which

$$\mid XY^{(n-h+1)} \mid \leqq \frac{1}{\sigma^{(h)}} = \tau^{(n-h+1)} \qquad (h = 1, 2, \cdots, n).$$

By choosing X such that there is equality in one of these conditions, but that all other scalar products $XY^{(n-h+1)}$ vanish, the assertion follows after replacing h by $n - h + 1$.

then[14]

$$G(Z^{(h)}) \geqq G(Y^{(h)}) = \tau^{(h)}.$$

Consider the $n + 1$ vectors

$$X^{(1)}, X^{(2)}, \cdots, X^{(n-h+1)}, \qquad Z^{(1)}, Z^{(2)}, \cdots, Z^{(h)}.$$

At most n of these are $\mathfrak{R}$-independent; hence the scalar products

$$X^{(i)}Z^{(j)} \qquad \begin{pmatrix} i = 1, 2, \cdots, n - h + 1 \\ j = 1, 2, \cdots, \quad h \end{pmatrix}$$

do not all vanish simultaneously, and at least one of them, say $X^{(i)}Z^{(j)}$, is different from zero. Since it is an element of $\mathfrak{T}$, therefore

$$| X^{(i)}Z^{(j)} | \geqq 1.$$

Now by (17)

$$| XY | \leqq F(X)G(Y),$$

for all points X and Y. Therefore

$$1 \leqq | X^{(i)} Z^{(j)} | \leqq F(X^{(i)})G(Z^{(j)}) \leqq F(X^{(n-h+1)})G(Z^{(h)}) = \frac{1}{\tau^{(h)}} G(Z^{(h)}),$$

as was to be proved.

From (28) and (29) in particular

(34) $$\sigma^{(1)} \leqq \left(\frac{\tau^{(1)}}{V}\right)^{1/n-1} \quad \text{and} \quad \tau^{(1)} \leqq (\sigma^{(1)}V)^{1/n-1},$$

so that if the minimum of $F(X)$ in $\mathfrak{T}$ is small, then the same is true for that of $G(Y)$, and vice versa.

11. The relation between the homogeneous and the inhomogeneous problem. The reciprocity formulae of the preceding paragraph can be applied to inhomogeneous problems. Let P be an arbitrary point in P_n which is not necessarily a lattice point; it can be written as

$$P = p_1X^{(1)} + \cdots + p_nX^{(n)}$$

where the p's lie in $\mathfrak{R}$. Put

$$p_h = -x_h + r_h \qquad (h = 1, 2, \cdots, n),$$

where x_h is an element of $\mathfrak{T}$ and

$$|r_h| \leqq \frac{1}{e} \qquad (h = 1, 2, \cdots, n).$$

[14] The minima $\sigma^{(h)}$ of $F(X)$ have the analogous property.

Then the lattice point $X = (x_1, \cdots, x_n)$ satisfies the inequality

$$F(P + X) = F\left(\sum_{h=1}^{n} r_h X^{(h)}\right) \leqq \frac{\sigma^{(n)}}{e},$$

or by (28)

$$F(P + X) \leqq \frac{1}{e\tau^{(1)}}. \tag{35}$$

This inequality cannot in general be improved, since

$$F\left(\frac{1}{z} X^{(n)} + X\right) \geqq \frac{1}{e\tau^{(1)}} \tag{36}$$

for all lattice points X, as follows immediately from the $\mathfrak{R}$-independence of the n vectors

$$X^{(1)}, X^{(2)}, \cdots, X^{(n-1)}, X^{(n)} + zX.$$

These two inequalities (35) *and* (36) *relate the inhomogeneous F-problem to the homogeneous G-problem,* in analogy with similar relations in many parts of mathematics.

As an application, consider the two polar distance functions

$$F(X) = \max(|\alpha_1 x_n - x_1|, \cdots, |\alpha_{n-1}x_n - x_{n-1}|, e^{-t}|x_n|),$$

$$G(Y) = \max(|y_1|, \cdots, |y_{n-1}|, e^t|\alpha_1 y_1 + \cdots + \alpha_{n-1}y_{n-1} + y_n|),$$

where t is a positive integer. Assume that the numbers $1, \alpha_1, \cdots, \alpha_{n-1}$ are $\mathfrak{R}$-independent, so that for all lattice points $Y = (y_1, \cdots, y_n) \neq 0$

$$\alpha_1 y_1 + \cdots + \alpha_{n-1}y_{n-1} + y_n \neq 0.$$

Then, as $t \to \infty$, the first minimum $\tau^{(1)}$ of $G(Y)$

$$\tau^{(1)} \to \infty.$$

Hence by (35), *for every* $\epsilon > 0$ *and for every point* $P = (p_1, \cdots, p_n)$ *there is a lattice point* $X = (x_1, \cdots, x_n)$ *satisfying the inequalities*

$$|\alpha_1 x_n - x_1 + p_1| < \epsilon, \cdots, |\alpha_{n-1}x_n - x_{n-1} + p_{n-1}| < \epsilon.$$

Thus we have established a result analogous to *Kronecker's* theorem.

12. A property of matrices. *Let*

$$\Omega = (a_{hk})_{h,k=1,2,\cdots,n}$$

be a matrix in $\mathfrak{R}$ *with determinant* 1; *then there is a matrix*

$$U = (u_{hk})_{h,k=1,2,\cdots,n}$$

with elements in $\mathfrak{T}$ *and determinant* 1, *such that the product matrix*

$$\Omega U = \Omega^* = (a^*_{hk})_{h,k=1,2,\cdots,n}$$

satisfies the equation

$$\prod_{h=1}^{n} \max_{k=1,2,\cdots,n} (|a^*_{hk}|) = 1.$$

Proof:[15] To the convex body $C(1)$ belonging to the distance function

$$F(X) = \max_{h=1,2,\cdots,n} \left(\left| \sum_{k=1}^{n} a_{hk} x_k \right| \right),$$

there are n lattice points $X^{(1)}, X^{(2)}, \cdots, X^{(n)}$ of determinant $D = 1$, such that the n minima

$$F(X^{(h)}) = \sigma^{(h)} \qquad (h = 1, 2, \cdots, n)$$

satisfy

$$0 < \sigma^{(1)} \leqq \sigma^{(2)} \leqq \cdots \leqq \sigma^{(n)}, \qquad \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)} = 1.$$

Let $X^{(k)} = (x_1^{(k)}, \cdots, x_n^{(k)})$, and X be the matrix

$$\mathrm{X} = (x_h^{(k)})_{h,k=1,2,\cdots,n}$$

with elements in $\mathfrak{T}$ and determinant 1. We introduce new coordinates $y_1, \cdots, y_n$ by putting

$$X = y_1 X^{(1)} + \cdots + y_n X^{(n)}, \text{ i.e., } x_h = \sum_{k=1}^{n} x_h^{(k)} x_k \qquad (h = 1, 2, \cdots, n);$$

then $F(X)$ changes into

$$F(X) = F'(Y) = \max_{h=1,2,\cdots,n} \left(\left| \sum_{k=1}^{n} a'_{hk} y_k \right| \right),$$

where

$$\Omega' = (a'_{hk})_{h,k=1,2,\cdots,n} = \Omega \mathrm{X}.$$

The n points $X = X^{(h)}$ are transformed into $Y = E^{(h)}$ $(h = 1, 2, \cdots, n)$; hence

$$F'(E^{(h)}) = \sigma^{(h)} \qquad (h = 1, 2, \cdots, n),$$

that is

$$\max_{h=1,2,\cdots,n} (|a'_{hk}|) = \sigma^{(k)} \qquad (k = 1, 2, \cdots, n). \tag{37}$$

[15] An analogous theorem in the real field was proved some time ago by C. L. Siegel in a letter to L. J. Mordell. The present proof and theorem, though not stated in Siegel's paper, are obtained from it with only slight changes by making use of the results in §9.

Hence every minor Δ_m of order m formed from the m first columns and m arbitrary rows of Ω' satisfies the inequality

$$(38) \qquad | \Delta | \leqq \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(m)}.$$

On the other hand, any determinant Δ of order m can be written as

$$\Delta = \sum_{h=1}^{n} a_h \delta_h,$$

where the a_h are the elements of its last column, and the δ_h their cofactors; therefore

$$\max_{h=1,2,\cdots,n} (| \delta_h |) \geqq | \Delta | \{ \max_{h=1,2,\cdots,n} (| a_h |)\}^{-1}.$$

We apply this inequality repeatedly to the determinant

$$\Delta_n = 1 = \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n)}$$

of Ω' and use (37) and (38); then it follows that *there exists*

an $(n-1)^{\text{th}}$ *order minor* Δ_{n-1} *of* Δ_n *formed from the* $n-1$ *first columns of* Ω' *and satisfying*

$$| \Delta_{n-1} | = \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n-1)};$$

an $(n-2)^{\text{th}}$ *order minor* Δ_{n-2} *of* Δ_{n-1} *formed from the* $n-2$ *first columns of* Ω' *and satisfying*

$$| \Delta_{n-2} | = \sigma^{(1)}\sigma^{(2)} \cdots \sigma^{(n-2)};$$

etc.; *a second order minor* Δ_2 *of* Δ_3 *formed from the two first columns of* Ω' *and satisfying*

$$| \Delta_2 | = \sigma^{(1)}\sigma^{(2)};$$

and finally an element Δ_1 *of* Δ_2 *lying in the first column of* Ω' *and satisfying*

$$| \Delta_1 | = \sigma^{(1)}.$$

Without loss of generality, we may assume that the determinants so constructed are exactly the principle determinants

$$\Delta_r = | a'_{hk} |_{h,k=1,2,\cdots,r} \qquad (r = 1, 2, \cdots, n).$$

We shall now construct a set of matrices of order n

$$U_m = \begin{pmatrix} 1 & 0 & \cdots & 0 & g_1^{(m)} & 0 & \cdots & 0 \\ & 1 & \cdots & 0 & g_2^{(m)} & 0 & \cdots & 0 \\ & & \ddots & \vdots & \vdots & \vdots & & \vdots \\ & & & 1 & g_{m-1}^{(m)} & 0 & \cdots & 0 \\ & & & & 1 & 0 & \cdots & 0 \\ & O & & & & 1 & \cdots & 0 \\ & & & & & & \ddots & \vdots \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\\1\\1\end{matrix}}\right\} m \text{ rows} \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n-m \text{ rows} \end{matrix} \qquad (m = 1, 2, \cdots, n),$$

where the g's lie in $\mathfrak{T}$, and U_1 is the unit matrix. If

$$\Omega_m = \Omega' U_1 U_2 \cdots U_m = (a_{hk}^{(m)})_{h,k=1,2,\cdots,m} \qquad (m = 1, 2, \cdots, n),$$

then $\Omega_1 = \Omega'$, and for $h, k = 1, 2, \cdots, n$

$$a_{hk}^{(m)} = a_{hk}^{(m-1)} \text{ if } k \neq m, \text{ and } a_{hm}^{(m)} = g_1^{(m)} a_{h1}^{(m-1)} + \cdots + g_{m-1}^{(m)} a_{hm-1}^{(m-1)} + a_{hm}^{(m-1)}.$$

The n principal determinants of Ω_m :

$$\Delta_r = | a_{hk}^{(m)} |_{h,k=1,2,\cdots,r} \qquad (r = 1, 2, \cdots, n)$$

are therefore equal to the corresponding ones of Ω_{m-1} and so of Ω'.

By construction, the elements of Ω_1 satisfy the inequalities

$$| a_{hk}^{(1)} | \leqq \sigma^{(k)} \qquad (h, k = 1, 2, \cdots, n),$$

and therefore also the inequalities

$$| a_{h1}^{(1)} | \leqq \sigma^{(h)} \qquad (h = 1, 2, \cdots, n).$$

Assume now that $U_1, \cdots, U_{m-1}$ were determined such that

$$(39) \qquad \begin{aligned} &| a_{hk}^{(m-1)} | \leqq \sigma^{(k)} && (h, k = 1, 2, \cdots, n); \\ &| a_{hk}^{(m-1)} | \leqq \sigma^{(h)} && \text{for } h = 1, 2, \cdots, n; k = 1, 2, \cdots, m-1. \end{aligned}$$

Then U_m, as we shall prove now, can be constructed such that Ω_m satisfies the stronger inequalities

$$(40) \qquad \begin{aligned} &| a_{hk}^{(m)} | \leqq \sigma^{(k)} && (h, k = 1, 2, \cdots, n); \\ &| a_{hk}^{(m)} | \leqq \sigma^{(h)} && \text{for } h = 1, 2, \cdots, n; k = 1, 2, \cdots, m. \end{aligned}$$

To this purpose put

$$a_{h1}^{(m-1)} \gamma_1 + \cdots + a_{hm-1}^{(m-1)} \gamma_{m-1} + a_{hm}^{(m-1)} = t_h(\gamma_1, \cdots, \gamma_{m-1}) = t_h \quad (h = 1, 2, \cdots, n),$$

and determine elements $\gamma_1, \gamma_2, \cdots, \gamma_{m-1}$ of $\mathfrak{K}$ such that

$$t_1 = t_2 = \cdots = t_{m-1} = 0.$$

This system of linear equations has the determinant Δ_{m-1}. On solving,

$$\Delta_{m-1}\gamma_r = \mp \Delta_{m-1,r} \qquad (r = 1, 2, \cdots, m-1),$$

where $\Delta_{m-1,r}$ is the $(m-1)^{\text{th}}$ order minor of Δ_m obtained by omitting the m^{th} row and the r^{th} column. Hence from (37),

$$| \gamma_r | = \left| \frac{\Delta_{m-1,r}}{\Delta_{m-1}} \right| \leqq \frac{\sigma^{(1)} \cdots \sigma^{(m)}}{\sigma^{(r)}} : (\sigma^{(1)} \cdots \sigma^{(m-1)}) = \frac{\sigma^{(m)}}{\sigma^{(r)}} \geqq 1.$$

Let the element $g_r^{(m)}$ of U_m now be the number in $\mathfrak{T}$ satisfying the inequality

$$| g_r^{(m)} - \gamma_r | < 1 \qquad (r = 1, 2, \cdots, m-1),$$

so that

$$| g_r^{(m)} | = \frac{\sigma^{(m)}}{\sigma^{(r)}}.$$

Then from the first system of inequalities (39) for $h = 1, 2, \cdots, n$

$$| a_{hm}^{(m)} | = | t_h(g_1^{(m)}, \cdots, g_{m-1}^{(m)}) | = | g_1^{(m)} a_{h1}^{(m-1)} + \cdots + g_{m-1}^{(m)} a_{hm-1}^{(m-1)} + a_{hm}^{(m-1)} |$$
$$\leqq \max \left(\frac{\sigma^{(m)}}{\sigma^{(1)}} \cdot \sigma^{(1)}, \cdots, \frac{\sigma^{(m)}}{\sigma^{(m-1)}} \cdot \sigma^{(m-1)}, \sigma^{(m)} \right) = \sigma^{(m)},$$

and from the second system for $h = 1, 2, \cdots, m$

$$| a_{hm}^{(m)} | = | t_h(g_1^{(m)}, \cdots, g_{m-1}^{(m)}) |$$
$$= | (g_1^{(m)} - \gamma_1) a_{h1}^{(m-1)} + \cdots + (g_{m-1}^{(m)} - \gamma_{m-1}) a_{hm-1}^{(m-1)} | < 1 \cdot \sigma^{(h)} = \sigma^{(h)}.$$

Since the remaining inequalities (40) are contained in (39), the matrix U_m has the required property. Hence if

$$U = XU_1U_2 \cdots U_n,$$

then this matrix satisfies the statement of our theorem.

13. A property of the product of n inhomogeneous linear polynomials in n variables. Let $\Omega = (a_{hk})_{h,k=1,2,\cdots,n}$ be again a matrix with elements in $\mathfrak{R}$ of determinant 1. We form the distance function

$$F(X | f) = \max_{h=1,2,\cdots,n} (e^{f_h} | a_{h1}x_1 + a_{h2}x_2 + \cdots + a_{hn}x_n |),$$

where $f_1, f_2, \cdots, f_n$ are n integers such that $f_1 + \cdots + f_n = 0$. By the theorem of last paragraph, there is a matrix U with elements in $\mathfrak{T}$ and determinant 1, such that the product matrix

$$\Omega^* = \Omega U = (a_{hk}^*)$$

satisfies the equation

$$\prod_{h=1}^{n} \max_{k=1,2,\cdots,n} (| a_{hk}^* |) = 1.$$

Let us choose the integers f_h^0 such that

$$e^{-f_h^0} = \max_{k=1,2,\cdots,n} (| a_{hk}^* |) \qquad (h = 1, 2, \cdots, n) \tag{41}$$

and put

$$a_{hk}^{**} = z^{f_h^0} a_{hk}^* \qquad (h, k = 1, 2, \cdots, n).$$

Then by the transformation $X = UY$, $F(X) | f^0)$ changes into a new distance function

$$F(X | f^0) = F'(Y) = \max_{h=1,2,\cdots,n} (| a_{h1}^{**} y_1 + \cdots + a_{hn}^{**} y_n |),$$

where now all coefficients a_{hk}^{**} satisfy the inequalities $|a_{hk}| \leqq 1$, and their determinant is still 1. Obviously, for all n $\mathfrak{K}$-independent vectors $Y^{(1)} = E^{(1)}$, $Y^{(2)} = E^{(2)}, \cdots, Y^{(n)} = E^{(n)}$, the value of this function

$$F'(Y^{(h)}) \leqq 1 \qquad (h = 1, 2, \cdots, n).$$

Therefore by the equation (25), necessarily

$$F'(Y^{(1)}) = F'(Y^{(2)}) = \cdots = F'(Y^{(n)}) = 1,$$

and so all minima of $F(X \mid f^0)$, where the f^0's are given by (41), have the same value 1, and in particular, the first minimum of $F(X \mid f^0)$ has the exact value $\frac{1}{\sqrt[n]{V}}$, where $V = 1$ is the volume of $F(X \mid f^0) \leqq 1$.

As an application, let $a_1, a_2, \cdots, a_n$ be any n elements of $\mathfrak{K}$, and $\eta_1, \eta_2, \cdots, \eta_n$ n elements of $\mathfrak{K}$ satisfying the equations

$$a_{h1}^{*}\eta_1 + \cdots + a_{hn}^{*}\eta_h + a_h = 0 \qquad (h = 1, 2, \cdots, n).$$

If $y_1, y_2, \cdots, y_n$ are the elements of $\mathfrak{T}$ for which

$$|y_h - \eta_h| \leqq \frac{1}{e} \qquad (h = 1, 2, \cdots, n),$$

then obviously

$$|a_{h1}^{*}y_1 + \cdots + a_{hn}^{*}y_n + a_h| \leqq e^{-f_h^0-1} \qquad (h = 1, 2, \cdots, n).$$

Hence the lattice point $X = (x_1, x_2, \cdots, x_n) = U^I Y$ *satisfies the inequalities*

$$|a_{h1}x_1 + \cdots + a_{hn}x_n + a_h| \leqq e^{-f_h^0-1} \qquad (h = 1, 2, \cdots, n),$$

and therefore the inequality

$$\prod_{h=1}^{n} |a_{h1}x_1 + \cdots + a_{hn}x_n + a_h| \leqq e^{-n}.$$

Here the constant e^{-n} on the right-hand side is the best possible, as is clear if, e.g. Ω is the unit matrix and all $a_h = 1/z$.

14. Distance functions in $\mathfrak{K}_\mathfrak{p}$. The field $\mathfrak{R}$ of all rational functions with coefficients in $\mathfrak{k}$ has valuations different from the "infinite" valuation $|x|$, which expresses the behavior of x at the point $z = \infty$.

Let ζ be any element of $\mathfrak{k}$, and $\mathfrak{p}$ the "finite" point $z = \zeta$. Then we define a valuation $|x|_\mathfrak{p}$ by putting for $x \neq 0$

$$|x|_\mathfrak{p} = e^{-f_\mathfrak{p}},$$

where $f_\mathfrak{p}$ is that integer, for which neither the numerator nor the denominator of the simplified fraction $(z - \zeta)^{-f_\mathfrak{p}}x$ are divisible by $z - \zeta$; we denote by $\mathfrak{K}_\mathfrak{p}$

the perfect extension of $\mathfrak{R}$ with respect to this valuation; it consists of all formal Laurent series

$$x = \alpha_f(z-\zeta)^f + \alpha_{f+1}(z-\zeta)^{f+1} + \alpha_{f+2}(z-\zeta)^{f+2} + \cdots$$

with coefficients in $\mathfrak{k}$, and if $\alpha_f \neq 0$, then $| x |_\mathfrak{p} = e^{-f}$.

Let now $F(X)$ be any special distance function of $\mathfrak{K}$; we use it as the measure for the size of X. Further let $F(X \mid \mathfrak{p})$ be a general distance function of $\mathfrak{K}_f$. Since

$$F((z-\zeta)^f X \mid \mathfrak{p}) = e^{-f} F(X \mid \mathfrak{p}),$$

this distance function may assume arbitrarily small values, if X lies in the modul Λ_n of all lattice points. By (5), there is a constant $\Gamma_\mathfrak{p} > 0$ such that

$$F(X \mid \mathfrak{p}) \leqq \Gamma_\mathfrak{p} \, | X |_\mathfrak{p} ;$$

here for $X = (x_1, \cdots, x_n)$

$$| X |_\mathfrak{p} = \max (| x_1 |_\mathfrak{p}, \cdots, | x_n |_\mathfrak{p}).$$

Hence

$$F(X \mid \mathfrak{p}) \leqq \Gamma_\mathfrak{p} \text{ for all lattice points } X.$$

Let t be an integer such that

$$e^{-t} \leqq \Gamma_\mathfrak{p}, \quad \text{i.e.} \quad t \geqq \log\left(\frac{1}{\Gamma_\mathfrak{p}}\right),$$

and $C(e^{-t} \mid \mathfrak{p})$ the convex set of all points X in P_n for which

$$F(X \mid \mathfrak{p}) \leqq e^{-t}.$$

Then the set $\mathfrak{m}(-t \mid \mathfrak{p})$ of all lattice points in $C(e^{-t} \mid \mathfrak{p})$ contains with X and Y also $aX + bY$, when a and b lie in $\mathfrak{T}$; it is therefore an $\mathfrak{T}$-modul. By the general theory of polynomial ideals,[16] this modul has a basis of n lattice points

$$P^{(k)} = (p_1^{(k)}, \cdots, p_n^{(k)}) \qquad (k = 1, 2, \cdots, n),$$

such that every point X in Λ_n belongs to $\mathfrak{m}(-t \mid \mathfrak{p})$, if and only it can be written as

$$X = y_1 P^{(1)} + \cdots + y_n P^{(n)} \quad \text{with } y_1, \cdots, y_n \text{ in } \mathfrak{T}.$$

The determinant

$$D(-t) = | p_h^{(k)} |_{h,k=1,2,\cdots,n} \neq 0,$$

and therefore the number

$$\Delta(-t) = | D(-t) |$$

is positive.

---

[16] Compare the basis theorem in §80 of van der Waerden's "Moderne Algebra", Vol. II, 1st ed.

The function $F(X)$ changes into a new distance function

$$F'(Y) = F(X) = F(\Omega Y), \qquad \Omega = (p_h^{(k)})_{h,k=1,2,\cdots,n})$$

by the transformation (42). The convex body $F'(Y) \leqq 1$ has the volume

$$V' = \Delta(-t)^{-1} V,$$

where V denotes the volume of $F(X) \leqq 1$. By the results in §9, there are n lattice points $Y^{(1)}, \cdots, Y^{(n)}$ with determinant 1, such that

$$F'(Y^{(1)}) \cdots F(Y^{(n)}) = \frac{\Delta(-t)}{V}.$$

The transformed lattice points $X^{(1)}, \cdots, X^{(n)}$ given by

$$X^{(k)} = \Omega Y^{(k)} = (x_1^{(k)}, \cdots, x_n^{(k)}) \qquad (k = 1, 2, \cdots, n)$$

have the determinant

$$D(-t) = | x_h^{(k)} |_{h,k=1,2,\cdots,n},$$

and satisfy the relations

$$F(X^{(1)}) \cdots F(X^{(n)}) = \frac{\Delta(-t)}{V}, \qquad F(X^{(k)} \mid \mathfrak{p}) \leqq e^{-t} \qquad (k = 1, 2, \cdots, n).$$

It is not difficult to prove that for large t

$$\Delta(-t) = O(e^{nt}), \qquad | D(-t) |_{\mathfrak{p}} = O(e^{-t}).$$

In the following case, sharper results are obtained. Let

$$F(X \mid \mathfrak{p}) = \max_{h=1,2,\cdots,m} (| a_{h1}x_1 + \cdots + a_{hn-m}x_{n-m} + x_{n-m+h} |_{\mathfrak{p}}),$$

where the a's are elements in $\mathfrak{K}_{\mathfrak{p}}$ such that

$$| a_{hk} |_{\mathfrak{p}} \leqq 1 \qquad \begin{pmatrix} h = 1, 2, \cdots, m \\ k = 1, 2, \cdots, n \end{pmatrix}.$$

Then to every positive integer t there are elements A_{hk} in $\mathfrak{T}$ satisfying

$$| a_{hk} - A_{hk} |_{\mathfrak{p}} \leqq e^{-t} \qquad \begin{pmatrix} h = 1, 2, \cdots, m \\ k = 1, 2, \cdots, n \end{pmatrix}.$$

Hence, if $y_1, \cdots, y_n$ belong to $\mathfrak{T}$, and $x_1, \cdots, x_n$ are defined by

$$x_1 = y_1, \cdots, x_{n-m} = y_{n-m};$$

(42) $$x_{n-m+h} = (z - \zeta)^t y_{n-m+h} - (A_{h1}y_1 + \cdots + A_{hn-m}y_{n-m}),$$

$$(h = 1, 2, \cdots, m),$$

then $F(X \mid \mathfrak{p}) \leqq e^{-t}$. Let $F'(Y) = F(X)$ be the special distance function in $\mathfrak{K}$ derived from $F(X)$ by the transformation (42). Then $F'(Y) \leqq 1$ has the

volume $|(z - \zeta)^{-mt}| V = e^{-mt}V$. Hence there are n $\mathfrak{K}$-independent lattice points $Y^{(1)}, \cdots, Y^{(n)}$ of determinant 1 such that

$$F'(Y^{(1)}) \cdots F'Y^{(n)}) = \frac{e^{mt}}{V}.$$

The n lattice points $X^{(1)}, \cdots, X^{(n)}$ derived from these by (42) have the determinant $(z - \zeta)^{mt}$ and satisfy the conditions

$$F(X^{(1)}) \cdots F(X^{(n)}) = \frac{e^{mt}}{V}, \qquad F(X^{(k)} \mid \mathfrak{p}) \leqq e^{-t} \qquad (k = 1, 2, \cdots, n).$$

MANCHESTER, ENGLAND.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

CONCRETE REPRESENTATION OF ABSTRACT (*L*)-SPACES AND THE MEAN ERGODIC THEOREM

BY SHIZUO KAKUTANI

(Received February 6, 1940)

1. **Introduction.**[1] Let Ω be an abstract space where a completely additive measure is defined. As is well known, the totality of all the real-valued measurable functions $x(t)$ which are absolutely integrable on Ω constitutes a Banach space $L(\Omega)$ with $\| x \| = \int_\Omega | x(t) | \, dt$ as its norm. Although the space $L(\Omega)$ is not necessarily separable,[2] it may always be considered as a semi-ordered Banach space.[3] Indeed, if we denote, for any pair of elements $x(t)$ and $y(t) \in L(\Omega)$, by $x \geqq y$ (or $y \leqq x$) the relation that $x(t) \geqq y(t)$ almost everywhere on Ω, then the following conditions are satisfied ($x, y, z, w \in L(\Omega)$, λ = scalar):

(I) $x \geqq y$ *and* $y \geqq x$ *imply* $x = y$,[4]

(II) $x \geqq y$ *and* $y \geqq z$ *imply* $x \geqq z$,

(III) $x \geqq y$ *and* $\lambda \geqq 0$ *imply* $\lambda x \geqq \lambda y$,

(IV) $x \geqq y$ *implies* $x + z \geqq y + z$ *for any* z,

(V) $x_n \geqq y_n$, $x_n \to x$ (*strongly*) *and* $y_n \to y$ (*strongly*) *imply* $x \geqq y$,

(VI) *to any pair of elements* x *and* y, *there exists a maximum* $z = x \vee y$ *such that* $z \geqq x$, $z \geqq y$, *and* $z \leqq z'$ *for any* z' *with* $z' \geqq x$, $z' \geqq y$,

(VII) *to any pair of elements* x *and* y, *there exists a minimum* $w = x \wedge y$ *such that* $w \leqq x$, $w \leqq y$, *and* $w \geqq w'$ *for any* w' *with* $w' \leqq x$, $w' \leqq y$.

Moreover, this semi-ordered Banach space $L(\Omega)$ has the following important property:

(VIII) $x \geqq 0$ *and* $y \geqq 0$ *imply* $\| x + y \| = \| x \| + \| y \|$;

in orther words, calling x to be *positive* in case $x \geqq 0$, norm is additive on positive elements. Such a Banach space was introduced axiomatically by Garrett

---

[1] The principal results of this paper were previously announced in S. Kakutani [7]. In [7] we have tacitly assumed the condition (IX).

[2] There are two typical cases when $L(\Omega)$ is not separable. The first one is the case of the Haar's measure of a non-separable bicompact topological group, and the second one is the case of the linear measure in the plane. In the first case, the total space Ω is of finite measure and every measurable subset Ω' of Ω with $m(\Omega') > 0$ determines a non-separable Banach space $L(\Omega')$. In the second case, the total space is not expressible as a sum of a countable infinite number of subsets of finite measure, while $L(\Omega')$ is separable for every measurable subset Ω' of Ω with $m(\Omega') < \infty$.

[3] It is to be noted that in the first case (see footnote (2)) there exists an element $x_0 > 0$ (for example, a function $x_0(t)$ which is identically equal to 1) such that $x_0 \wedge x > 0$ for any $x > 0$, while there exists no such element in the second case.

[4] $x = y$ means that we have $x(t) = y(t)$ almost everywhere on Ω.

Birkhoff [3]. He has introduced the space of this type as a generalization of the concrete Banach space (L) (i.e., the space of all the real-valued measurable functions $x(t)$ which are absolutely integrable on $0 \leqq t \leqq 1$), and has discussed the iteration of bounded linear operations in such Banach spaces. We shall call a Banach space with the semi-ordering satisfying (I)–(VIII) an *abstract* (L)-*space* (notation: (AL)). Every Banach space $L(\Omega)$ is an (AL), and, in contrast to general abstract (L)-spaces, this will be called a *concrete* (L)-space.

In the present paper we shall discuss the converse problem, i.e., we shall investigate how it is possible to represent any abstract (L)-space (AL) by a concrete (L)-space $L(\Omega)$. In other words, given an abstract (L)-space (AL) with the semi-ordering satisfying (I)–(VIII), it is required to construct a space Ω and a completely additive measure defined on some Borel field of Ω such that the corresponding Banach space $L(\Omega)$ is equivalent (= isometric and lattice-isomorphic) to the given space (AL).

This problem is not always possible, if we have no further assumptions on (AL). In order to see this, we have only to notice that the property:

(IX) $x \wedge y = 0$ *implies* $\| x + y \| = \| x - y \|$,

which is always satisfied for any concrete (L)-space, does not necessarily follow from the conditions (I)–(VIII). Indeed, if we consider the (x, y)-plane with the usual semi-ordering: $(x_1, y_1) \geqq (x_2, y_2)$ if and only if $x_1 \geqq x_2$ and $y_1 \geqq y_2$ simultaneously, and define its norm by

$$\begin{aligned} \| (x, y) \| &= | x + y | \quad \text{if} \quad x \geqq 0, \ y \geqq 0 \quad \text{or} \quad x \leqq 0, \ y \leqq 0, \\ &= \sqrt{x^2 + y^2} \quad \text{if} \quad x \geqq 0, \ y \leqq 0 \quad \text{or} \quad x \leqq 0, \ y \geqq 0, \end{aligned}$$

then the conditions (I)–(VII) are all satisfied, and yet we have $\| (1, 0) + (0, 1) \| = \| (1, 1) \| = 2 > \| (1, 0) - (0, 1) \| = \| (1, -1) \| = \sqrt{2}$.

If, however, the conditions (I)–(IX) are all satisfied, then our problem has a solution. This will be proved in Theorem 7. The proof is divided into three parts (§§3, 4 and 5), and our principal idea is essentially contained in the papers of H. Freudenthal [4] and F. Wecken [14]. Moreover, it is to be noticed that every abstract (L)-space with the properties (I)–(VIII) can be provided with an equivalent norm which satisfies the additional condition (IX) (Theorem 1, §2).

In Theorem 9 (§6), we shall prove a mean ergodic theorem in abstract (L)-spaces. This is a generalization of a result of Garrett Birkhoff [3] and may be considered as one of the most general formulations of the mean ergodic theorem and Markoff's process. It is further to be noted that, by virtue of Theorem 1, the condition (IX) is unnecessary for the validity of this theorem.

In concluding the introduction, we shall list some elementary lemmas concerning the semi-ordered Banach space, which follow directly from the conditions (I)–(VII) and which are needed in the following discussions.

Lemma 1.1. $\lambda \geqq 0$ *implies* $\lambda(x \vee y) = \lambda x \vee \lambda y$, $\lambda(x \wedge y) = \lambda x \wedge \lambda y$.

Lemma 1.2. $(x \vee y) + z = (x + z) \vee (y + z)$, $(x \wedge y) + z = (x + z) \wedge (y + z)$.

Lemma 1.3. $(x \vee y) + (x \wedge y) = x + y$.

LEMMA 1.4. *$x_i \wedge x_j = 0$ $(i \neq j,\ i, j = 1, 2, \cdots, n)$ imply $x_1 + x_2 + \cdots + x_n = x_1 \vee x_2 \vee \cdots \vee x_n$.*

LEMMA 1.5. *$x = x_+ - x_-$, where $x_+ = x \vee 0$, $x_- = (-x) \vee 0$ and $x_+ \wedge x_- = 0$.*

LEMMA 1.6. *$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$, $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$.*

These lemmas will be found in H. Freudenthal [4] and L. Kantorovitch [8].

2. Change of norm.

THEOREM 1. *Every abstract (L)-space (AL) with the semi-ordering satisfying the conditions (I)–(VIII) can be provided with an equivalent norm which satisfies the conditions (I)–(IX).*

PROOF: Put $\| x \|^* = \| x_+ \| + \| x_- \|$ for any $x \epsilon (AL)$. Then we have $\| x \| \leqq \| x \|^*$, $\| \lambda x \|^* = | \lambda | \cdot \| x \|^*$ and $\| x + y \|^* \leqq \| x \|^* + \| y \|^*$ for any $x, y \epsilon (AL)$ and $\lambda \gtreqless 0$. The first two relations are almost trivial and the third one may be proved as follows: $\| x + y \|^* = \| (x + y)_+ \| + \| (x + y)_- \| \leqq \| x_+ + y_+ \| + \| x_- + y_- \|$ (since $0 \leqq (x + y)_+ \leqq x_+ + y_+$, $0 \leqq (x + y)_- \leqq x_- + y_-$) $= \| x_+ \| + \| y_+ \| + \| x_- \| + \| y_- \| = \| x \|^* + \| y \|^*$. Thus $\| x \|^*$ may be considered as a norm on (AL). Moreover, as is easily seen, the conditions (V), (VIII) and (IX) are all satisfied for this new norm $\| x \|^*$. Hence all what we have to prove is that the two norms $\| x \|$ and $\| x \|^*$ are equivalent.

In order to show this, denote by $(AL)^*$ the space (AL) metrized by the new norm $\| x \|^*$. We have only to prove that $(AL)^*$ is complete. For, since we have $\| x \| \leqq \| x \|^*$ for any x, the identical transformation: $x \to x$ is a bounded linear transformation which maps $(AL)^*$ biuniquely on (AL). Consequently if $(AL)^*$ is complete, then by a theorem of S. Banach [1] (pp. 40–41), this mapping must be bicontinuous and there exists a constant C such that $\| x \|^* \leqq C \| x \|$ for any x.

Now, in order to prove the completeness of $(AL)^*$, let $\{x_n\}$ $(n = 1, 2, \cdots)$ be a fundamental sequence in $(AL)^*$: $\lim_{m,n\to\infty} \| x_m - x_n \|^* = 0$. We have to show that there exists an $\bar{x} \epsilon (AL)^*$ such that $\lim_{n\to\infty} \| x_n - \bar{x} \|^* = 0$. Without the loss of generality we may assume that we have $\| x_m - x_n \|^* \leqq 2^{-n}$ for $m \geqq n$. Since $\| x_m - x_n \| \leqq \| x_m - x_n \|^*$ for any m and n, $\{x_n\}$ $(n = 1, 2, \cdots)$ is also a fundamental sequence in (AL), and, by the completeness of (AL), there exists an $\bar{x} \epsilon (AL)$ such that $\lim_{n\to\infty} \| x_n - \bar{x} \| = 0$. We shall show that we have also $\lim_{n\to\infty} \| x_n - \bar{x} \|^* = 0$. For this purpose, put $\bar{x}_{n,p} = x_n \vee x_{n+1} \vee \cdots \vee x_{n+p}$ for $p = 0, 1, 2, \cdots$; $n = 1, 2, \cdots$ $(\bar{x}_{n,0} = x_n)$. Then we have $\bar{x}_{n,p} \leqq \bar{x}_{n,p+1} \leqq \bar{x}_{n,p} + (x_{n+p+1} - x_{n+p})_+$ and $\| \bar{x}_{n,p+1} - \bar{x}_{n,p} \| \leqq \| (x_{n+p+1} - x_{n+p})_+ \| \leqq \| x_{n+p+1} - x_{n+p} \|^* \leqq 2^{-(n+p)}$ for $p = 0, 1, 2, \cdots$; $n = 1, 2, \cdots$. Consequently, since $\sum_{p=0}^{\infty} \| \bar{x}_{n,p+1} - \bar{x}_{n,p} \| \leqq \sum_{p=0}^{\infty} 2^{-(n+p)} = 2^{-(n-1)}$, $\lim_{p\to\infty} \bar{x}_{n,p} \equiv \bar{x}_n$ (strongly) exists and this limit $\bar{x}_n$ clearly satisfies $x_m \leqq \bar{x}_n$ and $\| \bar{x}_n - x_m \| \leqq 2^{-(n-1)}$ for $m \geqq n$. Since $\lim_{m\to\infty} \| x_m - \bar{x} \| = 0$ we have $\bar{x} \leqq \bar{x}_n$ and $\| \bar{x}_n - \bar{x} \| \leqq 2^{-(n-1)}$

for $n = 1, 2, \cdots$. Consequently, $\| x_n - \bar{x} \|^* \leqq \| \bar{x}_n - x_n \|^* + \| \bar{x}_n - \bar{x} \|^* = \| \bar{x}_n - x_n \| + \| \bar{x}_n - \bar{x} \| \leqq 2 \cdot 2^{-(n-1)}$ for $n = 1, 2, \cdots$, and thus we have proved that $\lim_{n\to\infty} \| x_n - \bar{x} \|^* = 0$.

The proof of Theorem 1 is now completed.

Thus we have proved that we can introduce in every abstract (L)-space an equivalent norm which satisfies the conditions (I)–(IX). Hence we shall assume hereafter (in §§3, 4 and 5) that all the conditions (I)–(IX) are satisfied.

3. **Direct decomposition.** The principal result of this chapter is stated in Theorem 2. In the case of the space of functions of bounded variation,[5] Theorem 2 has previously been shown by F. Wecken [14].

We begin with elementary lemmas.

LEMMA 3.1. $\| x \vee y - x' \vee y \| \leqq \| x - x' \|$, $\| x \wedge y - x' \vee y \| \leqq \| x - x' \|$. *Consequently,* $x_n \to x$ *(strongly) implies* $x_n \vee y \to x \wedge y$ *(strongly) and* $x_n \wedge y \to x \wedge y$ *(strongly) for any* y.

PROOF. We shall prove only the first relation. $x \vee y = (x' + (x - x')) \vee y \leqq (x' + (x - x')_+) \vee (y + (x - x')_+) = x' \vee y + (x - x')_+$ implies $x \vee y -$
1.2 $x' \vee y \leqq (x - x')_+$ and consequently $(x \vee y - x' \vee y)_+ \leqq (x - x')_+$. Analogously, we have $(x \vee y - x' \vee y)_- \leqq (x - x')_-$. Consequently $\| x \vee y - x' \vee y \| = \| (x \vee y - x' \vee y)_+ \| + \| (x \vee y - x' \vee y)_- \| \leqq \| (x - x')_+ \| + \| (x - x')_+ \| = \| x - x' \|$ (by (IX)).

LEMMA 3.2. $0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_n \leqq x_{n+1} \leqq \cdots \leqq y$ *implies the existence of* $\lim_{n\to\infty} x_n \equiv x'$ *(strongly) with* $0 \leqq x' \leqq y$.

PROOF: For each n we have (by (VIII)) $\sum_{i=1}^{n-1} \| x_{i+1} - x_i \| = \| \sum_{i=1}^{n-1} (x_{i+1} - x_i) \| = \| x_n - x_1 \| \leqq \| y - x_1 \|$. Hence $\sum_{i=1}^{\infty} \| x_{i+1} - x_i \| < \infty$ and consequently $\lim_{n\to\infty} x_n = x'$ (strongly) exists and $0 \leqq x' \leqq y$ (by (V)).

LEMMA 3.3. *For any* $x \geqq 0$ *and* $y \geqq 0$, $\lim_{n\to\infty} (nx \wedge y) = P_x(y)$ *(strongly) exists and* $0 \leqq P_x(y) \leqq y$.

PROOF: Clear from Lemma 3.3.

LEMMA 3.4. *For any* $x \geqq 0$, $y \to P_x(y)$ *is a projection operator defined for all* $y \geqq 0$.[6]

(3.1) $$P_x(y + z) = P_x(y) + P_x(z),$$

(3.2) $$y \leqq z \text{ implies } P_x(y) \leqq P_x(z),$$

(3.3) $$P_x(\lambda y) = \lambda P_x(y) \text{ for any } \lambda \geqq 0,$$

(3.4) $\| P_x(y) \| \leqq \| y \|$. *More generally,* $\| P_x(y) - P_x(y') \| \leqq \| y - y' \|$. *Consequently,* $y_n \to y$ *(strongly) implies* $P_x(y_n) \to P_x(y)$ *(strongly),*

(3.5) $$P_x(P_x(y)) = P_x(y).$$

---

[5] This is, indeed, one of the most familiar examples of abstract (L)-spaces. In this case every principal ideal is separable. See also S. Banach and S. Mazur [2].

[6] In order to define $P_x(y)$ for any $y \in (AL)$, we have only to put $P_x(y) = P_x(y_+) - P_x(y_-)$ for any y. It is clear that we have (3.1), (3.3) (for any $\lambda \gtreqless 0$), (3.4) and (3.5) for any y.

PROOF: By Lemma 1.2, $(nx \wedge y) + (nx \wedge z) = (nx + (nx \wedge z)) \wedge (y + (nx \wedge z)) = 2nx \wedge (nx + z) \wedge (y + nx) \wedge (y + z)$. Hence $nx \wedge (y + z) \leqq (nx \wedge y) + (nx \wedge z) \leqq 2nx \wedge (y + z)$. Making $n \to \infty$ we have (3.1). (3.2) is clear since $y \leqq z$ implies $nx \wedge y \leqq nx \wedge z$ for $n = 1, 2, \cdots$. (3.3) is also clear since, by Lemma 1.1, $nx \wedge \lambda y = \lambda\left(\frac{n}{\lambda} x \wedge y\right)$ for $n = 1, 2, \cdots$. The first relation of (3.4) is again clear, and the second one is a direct consequence of the fact that we have $\| nx \wedge y - nx \wedge y' \| \leqq \| y - y' \|$ for $n = 1, 2, \cdots$ (by Lemma 3.1). Lastly, (3.5) follows from the relation that $nx \wedge P_x(y) = \lim_{m\to\infty} (nx \wedge (mx \wedge y))$ (strongly, by Lemma 3.1) $= nx \wedge y$ for $n = 1, 2, \cdots$.

LEMMA 3.5.

(3.6) $\qquad x \leqq x'$ *implies* $P_x(y) \leqq P_{x'}(y)$,

(3.7) $\qquad x \wedge x' = 0$ *implies* $P_x(y) \wedge P_{x'}(y) = 0$,

(3.8) $\qquad x \wedge x' = 0$ *implies* $P_{x+x'}(y) = P_x(y) + P_{x'}(y)$,

(3.9) $\quad x_n \leqq x$ $(n = 1, 2, \cdots)$ *and* $x_n \to x$ *(strongly) imply* $P_{x_n}(y) \to P_x(y)$ *(strongly).*[7]

PROOF: (3.6) is clear since $x \leqq x'$ implies $nx \wedge y \leqq nx' \wedge y$ for $n = 1, 2, \cdots$. (3.7) is also clear since $x \wedge x' = 0$ implies $(nx \wedge y) \wedge (nx' \wedge y) = 0$ for $n = 1, 2, \cdots$. (3.8) follows from the relation: $n(x + x') \wedge y = (nx + nx') \wedge y = (nx \vee nx') \wedge y$ (by Lemma 1.4) $= (nx \wedge y) \vee (nx' \wedge y)$ (by Lemma 1.6) $= (nx \wedge y) + (nx' \wedge y)$ (by Lemma 1.4, since $(nx \wedge y) \wedge (nx' \wedge y) = 0$) for $n = 1, 2, \cdots$. Lastly we shall prove (3.9): $x_n \leqq x$ implies $mx_n \wedge y \leqq P_{x_n}(y) \leqq P_x(y)$ (by (3.6)) and consequently $\| P_x(y) - P_{x_n}(y) \| \leqq \| P_x(y) - mx_n \wedge y \| \leqq \| P_x(y) - mx \wedge y \| + \| mx \wedge y - mx_n \wedge y \| \leqq \| P_x(y) - mx \wedge y \| + m \| x - x_n \|$ (by Lemma 3.1). Now, for any $\epsilon > 0$ take an m_0 so large that we have $\| P_x(y) - m_0 x \wedge y \| < \epsilon/2$ and then n_0 so large that we have $m_0 \| x - x_n \| < \epsilon/2$ for $n > n_0$. Then we have $\| P_x(y) - P_{x_n}(y) \| < \epsilon$ for $n > n_0$. Since $\epsilon > 0$ is arbitrary, we have $P_{x_n}(y) \to P_x(y)$ (strongly).

LEMMA 3.6. *For any $x \geqq 0$ and $y \geqq 0$, $P_x(y) = 0$ is equivalent to $x \wedge y = 0$.*

PROOF: Since we have always $P_x(y) \geqq x \wedge y$, $P_x(y) = 0$ implies $x \wedge y = 0$. Conversely, $x \wedge y = 0$ implies $nx \wedge y = 0$ for $n = 1, 2, \cdots$ and consequently $P_x(y) = 0$.

LEMMA 3.7. $x \wedge (y - P_x(y)) = 0$ *for any $x \geqq 0$ and $y \geqq 0$.*

PROOF: By (3.5) and (3.1), we have $P_x(y - P_x(y)) + P_x(y) = P_x(y - P_x(y)) + P_x(P_x(y)) = P_x(y)$. Hence $P_x(y - P_x(y)) = 0$ and, by Lemma 3.6, $x \wedge (y - P_x(y)) = 0$.

---

[7] It is worth noting that $x_n > x$ $(n = 1, 2, \cdots)$ and $x_n \to x$ (strongly) do not necessarily imply $P_{x_n}(y) \to P_x(y)$ (strongly). For example, put $x_n = \frac{1}{n} y$ for $n = 1, 2, \cdots$. Then we have $x_n \geqq 0$ $(n = 1, 2, \cdots)$, $x_n \to x = 0$ (strongly) and yet $P_{x_n}(y) = y$ does not tend to $P_x(y) = P_0(y) = 0$.

DEFINITION 1. For any $x \geqq 0$ and $y \geqq 0$, $x \succ y$ (or $y \prec x$) means that we have $y \wedge u = 0$ for any $u \geqq 0$ with $x \wedge u = 0$.

LEMMA 3.8. *$x \succ P_x(y)$ for any $x \geqq 0$ and $y \geqq 0$.*

PROOF: $x \wedge u = 0$ implies $(nx \wedge y) \wedge u = 0$ for $n = 1, 2, \cdots$, and making $n \to \infty$ we have $P_x(y) \wedge u = 0$.

LEMMA 3.9. *$x \succ y$ is equivalent to $P_x(y) = y$.*

PROOF: By Lemma 3.8, $P_x(y) = y$ implies $x \succ P_x(y) = y$. Conversely, since $x \wedge (y - P_x(y)) = 0$ by Lemma 3.7, $x \succ y$ implies $y - P_x(y) = y \wedge (y - P_x(y)) = 0$ and consequently $y = P_x(y)$.

DEFINITION 2. A set $\mathbf{I}$ of positive elements of (AL) is said to be an *ideal* if the following conditions are satisfied:

$$x \in \mathbf{I} \text{ and } y \in \mathbf{I} \text{ imply } x + y \in \mathbf{I}, \tag{3.10}$$

$$x \in \mathbf{I} \text{ and } y \prec x \text{ imply } y \in \mathbf{I}, \tag{3.11}$$

$$x_n \in \mathbf{I}\ (n = 1, 2, \cdots) \text{ and } x_n \to x \text{ (strongly) imply } x \in \mathbf{I}. \tag{3.12}$$

LEMMA 3.10. *For any $x \geqq 0$ the set of all $y \geqq 0$ which satisfy $y \prec x$ is an ideal.*

PROOF: We have only to prove the following three statements:

$$x \succ y \text{ and } x \succ z \text{ imply } x \succ y + z, \tag{3.13}$$

$$x \succ y \text{ and } y \succ z \text{ imply } x \succ z, \tag{3.14}$$

$$x \succ y_n\ (n = 1, 2, \cdots) \text{ and } y_n \to y \text{ (strongly) imply } x \succ y. \tag{3.15}$$

(3.13) is clear since $x \wedge u = 0$ implies $y \wedge u = 0$ and $z \wedge u = 0$, and consequently $0 \leqq (y + z) \wedge u \leqq (y + z) \wedge (u + z) \wedge (y + u) \wedge (u + u) = y \wedge u + z \wedge u$ (by Lemma 1.2) $= 0$. (3.14) is also clear since $x \wedge u = 0$ implies $y \wedge u = 0$ (since $x \succ y$) and this again implies $z \wedge u = 0$ (since $y \succ z$). Lastly, (3.15) follows from Lemma 3.1. For, $y_n \wedge u = 0$ $(n = 1, 2, \cdots)$ and $y_n \to y$ (strongly) imply $y \wedge u = \lim_{n\to\infty} (y_n \wedge u) = 0$ (by Lemma 3.1).

DEFINITION 3. The ideal obtained in Lemma 3.10 is called the *principal ideal* with *unit* x and is denoted by $[x]$. It is clear that $y \in [x]$ and $y > 0$ implies $y \wedge x > 0$. For each principal ideal, the unit is not unique. For example, we have $[x] = [\lambda x]$ for any $\lambda > 0$. More generally, any $y \geqq 0$ which satisfies $x \succ y$ and $y \succ x$ simultaneously has the property: $[x] = [y]$.

The totality of all the positive elements of (AL) constitutes itself an ideal. We shall call this ideal a unit ideal. The unit ideal is not necessarily principal, and we have

THEOREM 2. *The unit ideal is decomposed into a direct sum of a (not necessarily countable) number of principal ideals. More precisely, there exists a family of principal ideals $\{[x_\alpha]\}$ $(\alpha \in \mathfrak{M})$ such that $x_\alpha \wedge x_\beta = 0$ for any $\alpha \neq \beta$, and any $y > 0$ can be uniquely expressed in the form: $y = \sum_{n=1}^{\infty} y_{\alpha_n}$, $y_{\alpha_n} \in [x_{\alpha_n}]$, where $\{\alpha_n\}$ $(n = 1, 2, \cdots)$ is a countable sequence of indices from $\mathfrak{M}$ which depends on y such that $P_{x_{\alpha_n}}(y) = y_{\alpha_n}$ $(n = 1, 2, \cdots)$ and $P_{x_\alpha}(y) = 0$ for other $\alpha \in \mathfrak{M}$.*

Proof: Let all the positive elements of (AL) be arranged in the well-ordered sequence: $z_0, z_1, z_2, \cdots, z_\alpha, \cdots, \alpha < \varphi$. We shall define the transfinite sequence $\{x_\alpha\}$, $\alpha < \varphi$, by transfinite induction. Put $x_0 = z_0$, and assume that x_β is already defined for all $\beta < \alpha$. Then put $x_\alpha = z_\alpha$ if $z_\alpha \wedge x_\beta = 0$ for any $\beta < \alpha$, and $x_\alpha = 0$ otherwise (i.e., if there exists at least one $\beta < \alpha$ with $z_\alpha \wedge x_\beta > 0$). In this way, x_α can be defined for all $\alpha < \varphi$. It is clear that we have $x_\alpha \wedge x_\beta = 0$ for any $\alpha \neq \beta$, and that for any $x > 0$ there exists at least one $\alpha < \varphi$ with $x_\alpha \wedge x > 0$. We shall show that this sequence $\{x_\alpha\}$ ($\alpha < \varphi$, omitting those α with $x_\alpha = 0$, and this set of indices will be denoted by $\mathfrak{M}$) is the required one.

In order to prove this, consider, for any $y > 0$, the set $\{P_{x_\alpha}(y)\}$ $(\alpha \in \mathfrak{M})$. Since $0 \leqq P_{x_\alpha}(y) \leqq y$ for any $\alpha \in \mathfrak{M}$ and since $P_{x_\alpha}(y) \wedge P_{x_\beta}(y) = 0$ for any $\alpha \neq \beta$ (by (3.7)), we have $\sum_{i=1}^{n} P_{x_{\alpha_i}}(y) = P_{x_{\alpha_1}}(y) \vee P_{x_{\alpha_2}}(y) \vee \cdots \vee P_{x_{\alpha_n}}(y) \leqq y$ for any finite system of indices $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ from $\mathfrak{M}$ (by Lemma 1.4). Hence the set of indices $\alpha \in \mathfrak{M}$ with $P_{x_\alpha}(y) > 0$ is at most countable,[8] and if we denote these by $\{\alpha_n\}$ $(n = 1, 2, \cdots)$, then the strong limit $\lim_{n \to \infty} \sum_{i=1}^{n} P_{x_{\alpha_i}}(y) \equiv \sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y)$ exists and $\sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y) \leqq y$. We shall prove that this is an equality. Indeed, if we have $y' \equiv y - \sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y) > 0$, then there must exist at least one index $\alpha \in \mathfrak{M}$ with $x_\alpha \wedge y' > 0$. This is, however, a contradiction since we have $0 \leqq x_\alpha \wedge y' \leqq x_\alpha \wedge (y - P_{x_\alpha}(y)) = 0$ for any $\alpha \in \mathfrak{M}$ (by Lemma 3.6).

Thus we have proved that there exists a sequence of indices $\{\alpha_n\}$ $(n = 1, 2, \cdots)$ from $\mathfrak{M}$ such that $y = \sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y)$ and $P_{x_\alpha}(y) = 0$ for other $\alpha \in \mathfrak{M}$. In order to prove the uniqueness of this expression, let us assume that we have $y = \sum_{n=1}^{\infty} y_{\beta_n}$, $0 < y_{\beta_n} \in [x_{\beta_n}]$ $(n = 1, 2, \cdots)$. Then we have $y \geqq y_{\beta_n}$ and consequently $P_{x_{\beta_n}}(y) \geqq P_{x_{\beta_n}}(y_{\beta_n}) = y_{\beta_n} > 0$ (by Lemma 3.9) for $n = 1, 2, \cdots$. Hence $\{\beta_n\}$ $(n = 1, 2, \cdots)$ is a subsequence of $\{\alpha_n\}$ $(n = 1, 2, \cdots)$. We shall prove that the totality of $\{\beta_n\}$ $(n = 1, 2, \cdots)$ coincides with $\{\alpha_n\}$ $(n = 1, 2, \cdots)$ and that we have $P_{x_{\beta_n}}(y) = y_{\beta_n}$ for $n = 1, 2, \cdots$. For, otherwise, we should have $y = \sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y) > \sum_{n=1}^{\infty} P_{x_{\beta_n}}(y) \geqq \sum_{n=1}^{\infty} y_{\beta_n} = y$ or $y = \sum_{n=1}^{\infty} P_{x_{\alpha_n}}(y) \geqq \sum_{n=1}^{\infty} P_{x_{\beta_n}}(y) > \sum_{n=1}^{\infty} y_{\beta_n} = y$, which is clearly a contradiction.

This concludes the proof of Theorem 2.

Theorem 3. *In order that the unit ideal is principal, it is necessary and sufficient that $\mathfrak{M}$ is at most countable.*

Proof: If there exists a unit $\mathbf{1}$ such that $\mathbf{1} > x$ for any $x > 0$, or equivalently, $\mathbf{1} \wedge x > 0$ for any $x > 0$, then we have $x'_\alpha \equiv \mathbf{1} \wedge x_\alpha > 0$ for any $\alpha \in \mathfrak{M}$. Since $x'_\alpha \wedge x'_\beta = 0$ for any $\alpha \neq \beta$, we have $0 < x'_{\alpha_1} + x'_{\alpha_2} + \cdots + x'_{\alpha_n} = x'_{\alpha_1} \vee x'_{\alpha_2} \vee \cdots \vee x'_{\alpha_n} \leqq \mathbf{1}$ for any finite system of indices $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ from $\mathfrak{M}$ (by Lemma 1.4). Hence $\mathfrak{M}$ must be at most countable.[9] Conversely, if $\mathfrak{M}$ is

[8] We have only to notice that for each n there exists only a finite number of indices α such that $\| P_{x_\alpha}(y) \| > 1/n$.

[9] Analogously as in footnote 8. We have only to notice that for each n there exists only a finite number of indices α such that $\| x'_\alpha \| > 1/n$.

at most countable: $\mathfrak{M} = \{\alpha_1, \alpha_2, \cdots\}$, then $\mathbf{1} \equiv \sum_{n=1}^{\infty} 2^{-n} \| x_{\alpha_n} \|^{-1} \cdot x_{\alpha_n}$ will be a unit. Indeed, $x > 0$ implies the existence of an index α_n with $x_{\alpha_n} \wedge x > 0$ and consequently $\mathbf{1} \wedge x \geqq 2^{-n} \cdot \| x_{\alpha_n} \|^{-1} \cdot x_{\alpha_n} \wedge x > 0$.

THEOREM 4. *Every separable abstract (L)-space has a unit.*

PROOF: If we put $x'_\alpha = \| x_\alpha \|^{-1} \cdot x_\alpha$ for any $\alpha \in \mathfrak{M}$, then we have $\| x'_\alpha \| = 1$ and $x'_\alpha \wedge x'_\beta = 0$ for any $\alpha \neq \beta$. Hence (by (IX)) we have $\| x'_\alpha - x'_\beta \| = \| x'_\alpha + x'_\beta \| = \| x'_\alpha \| + \| x'_\beta \| = 2$ for any $\alpha \neq \beta$. Since (AL) is separable by assumption, $\mathfrak{M}$ must be at most countable and, by Theorem 3, the unit ideal of (AL) must be principal.

REMARK. This result may also be obtained directly as follows: Let $\{x_n\}$ $(n = 1, 2, \cdots)$ be a countable set which is dense in the positive part of (AL). If we put $\mathbf{1} \equiv \sum_{n=1}^{\infty} 2^{-n} \cdot \| x_n \|^{-1} \cdot x_n$, then $\mathbf{1}$ is a unit of (AL). Indeed, for any $x > 0$ there exists a subsequence $\{x_{n_\nu}\}$ $(\nu = 1, 2, \cdots)$ of $\{x_n\}$ $(n = 1, 2, \cdots)$ such that $x_{n_\nu} \to x$ (strongly) and $x_{n_\nu} \wedge x \to x \wedge x$ (strongly) $= x > 0$. Hence $x_{n_\nu} \wedge x > 0$ for some ν, and this implies $\mathbf{1} \wedge x \geqq 2^{-n_\nu} \cdot \| x_{n_\nu} \|^{-1} \cdot x_{n_\nu} \wedge x > 0$.

This result was also obtained by H. Freudenthal [4].

4. Integral representation.

In §3 we have obtained a direct decomposition of the positive part of (AL) into principal ideals. Consequently, our problem of concrete representation is reduced to the case of a principal ideal, i.e., the case when the unit element $\mathbf{1}$ exists.[10] In this chapter we shall show that *the positive part of an abstract (L)-space with unit may be represented by an integral in some abstract Boolean algebra with unit.* This may be considered as a generalization of a well-known result of O. Nikodym [9] (see also S. Saks [10]), and is essentially contained in the paper of H. Freudenthal [4]. The proof given below, however, has some interest.

Let us denote by $\mathbf{1}$ the unit element which we assume to exist (throughout this chapter). Without loss of generality, we may assume that $\| \mathbf{1} \| = 1$.

DEFINITION 4.[11] A positive element $e \geqq 0$ is said to be a *characteristic element* of (AL) if we have $e \wedge (\mathbf{1} - e) = 0$, or equivalently by Lemma 1.2, $2e \wedge \mathbf{1} = e$. The totality of all characteristic elements of (AL) will be denoted by $\mathbf{E}$. It is clear that $e \in \mathbf{E}$ implies $0 \leqq e \leqq \mathbf{1}$.

LEMMA 4.1. *$P_x(\mathbf{1}) \in \mathbf{E}$ for any $x \geqq 0$.*

PROOF: $2P_x(\mathbf{1}) \wedge \mathbf{1} = \lim_{n\to\infty} (2(nx \wedge \mathbf{1}) \wedge \mathbf{1})$ (strongly by Lemma 3.1) $= \lim_{n\to\infty} (2nx \wedge \mathbf{1}) = P_x(\mathbf{1})$.

LEMMA 4.2. *$e \in \mathbf{E}$ is equivalent to $P_e(\mathbf{1}) = e$.*

PROOF: By Lemma 4.1, $P_e(\mathbf{1}) = e$ implies $e \in \mathbf{E}$. Conversely, $e \in \mathbf{E}$ is equivalent to $2e \wedge \mathbf{1} = e$ by definition. And if $2^n e \wedge \mathbf{1} = e$, then $2^{n+1} e \wedge \mathbf{1} = 2(2^n e \wedge \mathbf{1}) \wedge \mathbf{1} = 2e \wedge \mathbf{1} = e$. Hence, $P_e(\mathbf{1}) = \lim_{n\to\infty} (2^n e \wedge \mathbf{1}) = e$.

LEMMA 4.3. *$e \in \mathbf{E}$ implies $\mathbf{1} - e \in \mathbf{E}$.* (Clear.)

---

[10] See the last lines of the proof of Theorem 7.

[11] Cf. H. Freudenthal [4].

LEMMA 4.4. *$e_1 \in \mathbf{E}$ and $e_2 \in \mathbf{E}$ imply $e_1 \vee e_2 \in \mathbf{E}$ and $e_1 \wedge e_2 \in \mathbf{E}$.*

PROOF: $2(e_1 \vee e_2) \wedge \mathbf{1} = (2e_1 \vee 2e_2) \wedge \mathbf{1} = (2e_1 \wedge \mathbf{1}) \vee (2e_2 \wedge \mathbf{1})$ (by Lemma 1.6) $= e_1 \vee e_2$. $2(e_1 \wedge e_2) \wedge \mathbf{1} = (2e_1 \wedge 2e_2) \wedge \mathbf{1} = (2e_1 \wedge \mathbf{1}) \wedge (2e_2 \wedge \mathbf{1}) = e_1 \wedge e_2$.

LEMMA 4.5. *$e_n \in \mathbf{E}$ $(n = 1, 2, \cdots)$ and $e_n \to e$ (strongly) imply $e \in \mathbf{E}$.*

PROOF: $2e \wedge \mathbf{1} = \lim_{n\to\infty} (2e_n \wedge \mathbf{1})$ (by Lemma 2.1) $= \lim_{n\to\infty} e_n = e$.

LEMMA 4.6. *$\mathbf{E}$ is a Boolean algebra with $e_1 \vee e_2$, $e_1 \wedge e_2$ and $\mathbf{1} - e$ as its fundamental operations, and is closed in (AL) in the strong topology.*

PROOF: Clear from Lemmas 1.6, 4.3, 4.4 and 4.5.

THEOREM 5. *For any $x \geqq 0$ there exists a system of characteristic elements $\{e(\lambda)\}$ $(0 \leqq \lambda < \infty)$, called the resolution of unity, such that*

$$\lambda \leqq \mu \text{ implies } e(\lambda) \leqq e(\mu), \tag{4.1}$$

$$\lambda_n \leqq \lambda\ (n = 1, 2, \cdots) \text{ and } \lambda_n \to \lambda \text{ imply } e(\lambda_n) \to e(\lambda) \text{ (strongly)}, \tag{4.2}$$

$$e(0) = 0, \lim_{\lambda\to\infty} e(\lambda) \equiv e(\infty) = \mathbf{1}, \tag{4.3}$$

$$e \in \mathbf{E} \text{ and } e \leqq e(\lambda) \text{ imply } P_e(x) \leqq \lambda e, \tag{4.4}$$

$$e \in \mathbf{E} \text{ and } e \leqq \mathbf{1} - e(\lambda) \text{ imply } P_e(x) \geqq \lambda e. \tag{4.5}$$

PROOF: Put $e(\lambda) = P_{(\lambda \mathbf{1} - x)_+}(\mathbf{1})$. Then $e(\lambda)$ is characteristic by Lemma 4.1. Since $(\lambda\mathbf{1} - x)_+ \leqq (\mu\mathbf{1} - x)_+$ for $\lambda \leqq \mu$, (4.1) is a direct consequence of (3.6). Analogously, (4.2) is a direct consequence of (3.9), if we observe that we have $(\lambda_n\mathbf{1} - x)_+ = (\lambda_n\mathbf{1} - x) \vee 0 \to (\lambda\mathbf{1} - x) \vee 0 = (\lambda\mathbf{1} - x)_+$ (strongly by Lemma 3.1). The first part of (4.3) is almost evident; for, we have $e(0) = P_{(-x)_+}(\mathbf{1}) = P_0(\mathbf{1}) = 0$. Before coming to the proof of the second part, we shall prove (4.4) and (4.5). $e \in \mathbf{E}$ and $e \leqq e(\lambda)$ imply $0 \leqq e \wedge (x - \lambda\mathbf{1})_+ \leqq e(\lambda) \wedge (x - \lambda\mathbf{1})_+ = \lim_{n\to\infty} (n(\lambda\mathbf{1} - x)_+ \wedge \mathbf{1} \wedge (x - \lambda\mathbf{1})_+)$ (strongly by Lemma 3.1) $= 0$ (by Lemma 1.5), and consequently $P_e((x - \lambda\mathbf{1})_+) = 0$. Hence the trivial relation $x \leqq (x - \lambda\mathbf{1})_+ + \lambda\mathbf{1}$ implies $P_e(x) \leqq P_e((x - \lambda\mathbf{1})_+) + P_e(\lambda\mathbf{1}) = P_e(\lambda\mathbf{1}) = \lambda e$ (by (3.2), (3.1), (3.3) and Lemma 4.2). In the same manner, $e \in \mathbf{E}$ and $e \leqq \mathbf{1} - e(\lambda)$ imply $0 \leqq e \wedge (\lambda\mathbf{1} - x)_+ \leqq (\mathbf{1} - e(\lambda)) \wedge (\lambda\mathbf{1} - x)_+ = (\mathbf{1} - P_{(\lambda\mathbf{1}-x)_+}(\mathbf{1})) \wedge (\lambda\mathbf{1} - x)_+ = 0$ (by Lemma 3.7), and consequently $P_e((\lambda\mathbf{1} - x)_+) = \lim_{n\to\infty} (ne \wedge (\lambda\mathbf{1} - x)_+)$ (strongly) $= 0$. Hence the trivial relation $x + (\lambda\mathbf{1} - x)_+ \geqq \lambda\mathbf{1}$ implies $P_e(x) = P_e(x) + P_e((\lambda\mathbf{1} - x)_+) \geqq P_e(\lambda\mathbf{1}) = \lambda e$. Thus (4.4) and (4.5) are proved.

Now, in order to prove the second relation of (4.3), put $\lim_{\lambda\to\infty} e(\lambda) = e(\infty)$ (strongly), which surely exists and belongs to $\mathbf{E}$ by Lemma 4.5. If we further put $e'(\infty) = \mathbf{1} - e(\infty)$, then $e'(\infty) \in \mathbf{E}$ by Lemma 4.3. Since $e'(\infty) \leqq \mathbf{1} - e(\lambda)$ for any $\lambda > 0$, we have $P_{e'(\infty)}(x) \geqq \lambda e'(\infty)$ and consequently $\| P_{e'(\infty)}(x) \| \geqq \lambda \| e'(\infty) \|$ for any $\lambda > 0$. From this follows directly that we have $e'(\infty) = 0$ and $e(\infty) = \mathbf{1}$.

Thus the proof of Theorem 5 is completed.

THEOREM 6. *Each positive element $x \geqq 0$ of (AL) can be expressed in the form:*

$$x = \int_0^\infty \lambda \, de(\lambda), \tag{4.6}$$

where the integration is of abstract Radon-Stieltjes type and $\{e(\lambda)\}$ $(0 \leqq \lambda < \infty)$ is the resolution of unity obtained in Theorem 5.

PROOF: For any division Δ: $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \Lambda < \infty$ of the interval $(0, \Lambda)$ with $0 < \lambda_i - \lambda_{i-1} < \epsilon$ $(i = 1, 2, \cdots, n)$, we have

$$m(\Delta) \equiv \sum_{i=1}^n \lambda_{i-1}(e(\lambda_i) - e(\lambda_{i-1})) \leqq \sum_{i=1}^n P_{(e(\lambda_i)-e(\lambda_{i-1}))}(x) = P_{e(\Lambda)}(x)$$
$$\leqq \sum_{i=1}^n \lambda_i(e(\lambda_i) - e(\lambda_{i-1})) \equiv M(\Delta)$$

and

$$M(\Delta) - m(\Delta) = \sum_{i=1}^n (\lambda_i - \lambda_{i-1})(e(\lambda_i) - e(\lambda_{i-1}))$$
$$\leqq \epsilon \sum_{i=1}^n (e(\lambda_i) - e(\lambda_{i-1})) = \epsilon(e(\Lambda) - e(0)) \leqq \epsilon \mathbf{1}.$$

These relations follow directly from the fact that we have $e(\lambda_i) - e(\lambda_{i-1}) \leqq e(\lambda_i)$ and $e(\lambda_i) - e(\lambda_{i-1}) \leqq \mathbf{1} - e(\lambda_{i-1})$ for $i = 1, 2, \cdots, n$. Hence we have (by making $\epsilon \to 0$) $P_{e(\Lambda)}(x) = \int_0^\Lambda \lambda \, de(\lambda)$, and, by making $\Lambda \to \infty$, we have the required relation (4.6) (since $\Lambda \to \infty$ implies $e(\Lambda) \to \mathbf{1}$ (strongly) and $P_{e(\Lambda)}(x) \to P_{\mathbf{1}}(x) = x$ (strongly) by (3.9)).

The proof of Theorem 6 is hereby completed.

REMARK. Theorems 5 and 6 are also valid even if there exists no unit element in (AL); for, we have only to consider the principal ideal $[x]$.

5. **Concrete representation.** In §4 we have seen that any positive element of an (AL) with unit can be represented by an integral in some abstract Boolean algebra. We shall show, in this chapter, that this abstract Boolean algebra (with a unit element) can be represented by a concrete one with a completely additive measure, and that the abstract integration can be substituted by a concrete one.

THEOREM 7. *To any abstract (L)-space (AL) satisfying the conditions* (I)–(IX), *with a unit element there corresponds a totally disconnected (bicompact) topological space Ω and a completely additive measure defined on a Borel field of Ω, such that (AL) is isometric and lattice-isomorphic to the Banach space $L(\Omega)$.*

PROOF: We shall first treat the case when the unit exists. By Lemma 4.6, the totality $\mathbf{E}$ of all the characteristic elements e of (AL) constitutes a Boolean algebra with $\mathbf{1}$ as its unit element. Hence, by a theorem of M. H. Stone [12]-

H. Wallman [13], $\mathbf{E}$ may be represented by a concrete Boolean algebra $\mathbf{K}$ of all the simultaneously open and closed subsets of a totally disconnected bicompact topological space Ω. Let E be the element of $\mathbf{K}$ which corresponds to the element e of $\mathbf{E}$. Then $m(E) = \| e \|$ is clearly a finitely additive measure defined on $\mathbf{K}$, and, if we put $d(E_1, E_2) = m(E_1 + E_2 - E_1E_2)$, then (denoting by e_1 and e_2 the corresponding elements in $\mathbf{E}$), $d(E_1, E_2) = m(E_1 + E_2 - E_1E_2) = m((E_1 + E_2) - E_2) + m((E_1 + E_2) - E_1) = \| e_1 \vee e_2 - e_2 \| + \| e_1 \vee e_2 - e_1 \| = \| (e_1 - e_2) \vee 0 \| + \| 0 \vee (e_2 - e_1) \| = \| e_1 - e_2 \|$ (by IX)); i.e., with this metric $d(E_1, E_2)$ $\mathbf{E}$ is isometric with $\mathbf{K}$.

We shall next prove that this $m(E)$ is completely additive on $\mathbf{K}$. For this purpose, it suffices to show that no element E of $\mathbf{K}$ can be expressed as a sum of a countable infinite number of non-vacuous disjoint sets $\{E_n\}$ $(n = 1, 2, \cdots)$ of $\mathbf{K}$. Indeed, if we have $E = \sum_{n=1}^{\infty} E_n$, $E \epsilon \mathbf{K}$, $E_n \epsilon \mathbf{K}$ $(n = 1, 2, \cdots)$ and $E_mE_n = 0$ $(m \neq n)$, then the closed set E is covered by a system of open sets $\{E_n\}$ $(n = 1, 2, \cdots)$. Since the space Ω is bicompact, E is covered by a finite number of E_n, and this is clearly impossible since each E_n $(n = 1, 2, \cdots)$ is non-vacuous and $E_m \cdot E_n = 0$ for $m \neq n$.

Thus we have proved that $m(E)$ is completely additive on $\mathbf{K}$. Hence, by a theorem of E. Hopf [5] (p. 2), $m(E)$ can be extended to the least Borel field $B(\mathbf{K})$ containing $\mathbf{K}$, and it will be easily seen that the residual class of $B(\mathbf{K})$ modulo the ideal of all the sets of measure zero of $B(\mathbf{K})$ forms a Boolean algebra which is isometric and lattice-isomorphic to $\mathbf{E}$.[12] Thus we have obtained a space Ω and a completely additive measure $m(E)$ defined on the Borel field $B(\mathbf{K})$ of subsets E of Ω which is isometric and lattice-isomorphic to $\mathbf{E}$ (if we neglect the sets of measure zero); and it is now an easy matter to show (under the condition (IX)) that (AL) is isometric and lattice-isomorphic to the Banach space $L(\Omega)$ which is determined by the measure just obtained above. Indeed, the integral representation obtained in Theorem 6 is simply the one-to-one norm-preserving correspondence of the positive part of (AL) and $L(\Omega)$, and from this follows (by virtue of (IX)) that (AL) and $L(\Omega)$ are isometric and lattice-isomorphic to each other.

Thus Theorem 7 is proved under the condition that the unit element exists. In order to discuss the general case, we have only to appeal to Theorem 2. Indeed, we have only to consider the family of spaces $\{\Omega_\alpha\}$, $\alpha \epsilon \mathfrak{M}$, corresponding to the principal ideals $\{[x_\alpha]\}$, $\alpha \epsilon \mathfrak{M}$, and to put $\Omega = \sum_{\alpha \epsilon \mathfrak{M}} \Omega_\alpha$. As is easily

[12] For this purpose we have only to show that for any $E \epsilon B(\mathbf{K})$ with $E = \sum_{n=1}^{\infty} E_n$, $E_n \epsilon \mathbf{K}$ $(n = 1, 2, \cdots)$ and $E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots$, there exists an $E' \epsilon \mathbf{K}$ such that $E' \supset E$ and $m(E' - E) = 0$. In order to show this, consider the corresponding elements $\{e_n\}$ $(n = 1, 2, \cdots)$ from $\mathbf{E}$. Then we have $e_1 \leqq e_2 \leqq \cdots \leqq e_n \leqq \cdots$, and $\| e_n \| = m(E_n) \leqq m(E) < \infty$. Hence $\lim_{n\to\infty} e_n = e'$ exists and, if we denote by E' the element of $\mathbf{K}$ which corresponds to e', then E' satisfies $E' \supset E_n$ $(n = 1, 2, \cdots)$ and $m(E') = \| e' \| = \lim_{n\to\infty} \| e_n \| = \lim_{n\to\infty} m(E_n)$. Consequently $E' \supset \sum_{n=1}^{\infty} E_n = E$ and $m(E' - E) = \lim m(E' - E_n) = 0$.

seen, $L(\Omega)$ is the totality of all the measurable functions $x(t)$ defined all over $\Omega = \sum_{\alpha \in \mathfrak{M}} \Omega_\alpha$, for which there exists at most a countable number of indices $\{\alpha_n\}$ $(n = 1, 2, \cdots)$ (depending on $x(t)$) such that we have $\sum_{n=1}^{\infty} \int_{\Omega_{\alpha_n}} | x(t) | \, dt < \infty$ and $x(t) = 0$ almost everywhere on other Ω_α. It is clear that (AL) and $L(\Omega)$ are isometric and lattice-isomorphic to each other.

The proof of Theorem 7 is now completed.

If, moreover, the space (AL) is separable, then $B(\mathbf{K})$ is also separable with respect to the metric $d(E_1, E_2) = m(E_1 + E_2 - E_1E_2)$. Furthermore, by Theorem 4, this (AL) has a unit $\mathbf{1}$ with $\| \mathbf{1} \| = 1$. Consequently $B(\mathbf{K})$ has a unit element $\mathbf{1}$ with $m(\mathbf{1}) = 1$. Hence, by a well-known argument, $B(\mathbf{K})$ can be embedded isometrically and lattice-isomorphically in a Boolean algebra of all the measurable sets of the unit interval $(0, 1)$.[13]

THEOREM 8. *Every separable abstract* (L)*-space* (*satisfying* (I)–(IX)) *can be embedded isometrically and lattice-isomorphically into the Banach space* (L) (= *the concrete* (L)*-space defined on the unit interval* $(0, 1)$ *with respect to the ordinary Lebesgue measure*).

6. **Mean ergodic theorem in abstract (L)-spaces.**

DEFINITION 5. Let T be a bounded linear operation which maps a semi-ordered Banach space into itself. T is said to be *positive* if we have $T(x) \geqq 0$ for any $x \geqq 0$.

THEOREM 9. *Let* T *be a positive bounded linear operation which maps an abstract* (L)*-space* (AL) (*satisfying the conditions* (I)–(VIII)) *into itself. If there exists a constant* C *such that* $\| T^n \| \leqq C$ *for* $n = 1, 2, \cdots$, *and if there exists for any* $x \geqq 0$ *an* $x_0 \geqq 0$ *such that* $x_n \equiv \frac{1}{n}(x + T(x) + \cdots T^{n-1}(x)) \leqq x_0$ *for* $n = 1, 2, \cdots$, *then for any* $x \in (AL)$ *the sequence* $\{x_n\}$ $(n = 1, 2, \cdots)$ *converges strongly to a point* $\bar{x} \in (AL)$, *i.e., mean ergodic theorem is valid in* (AL).

REMARK. This is a generalization of a result of Garrett Birkhoff [3]. He has assumed that T preserves the norm of positive elements (i.e., $\| T(x) \| = \| x \|$ for any $x \geqq 0$), and has only proved that $\{f(x_n)\}$ $(n = 1, 2, \cdots)$ converges for any bounded linear functional $f(x)$ defined on (AL). (Since the *weak completeness* of (AL) (see Theorem 11) was not proved by him, the weak convergence of $\{x_n\}$ $(n = 1, 2, \cdots)$ was not yet established).

PROOF OF THEOREM 9. By Theorem 1, we have only to discuss the case when the conditions (I)–(IX) are all satisfied. Hence we shall assume, appealing to Theorem 7, that (AL) is represented isometrically and lattice-isomorphically by a concrete (L)-space $L(\Omega)$. Ω may not be a sum of a countable infinite number of subsets of finite measure. By a theorem of K. Yosida [15] and the author [7] (see also F. Riesz [10]), we have only to prove that, for any $x \in L(\Omega)$, the se-

[13] Cf. S. Bochner and J. v. Neumann, *On compact solutions of operational-differential equations I*, Annals of Math., Vol. 36 (1935), p. 264, Footnote 17.

quence $\{x_n\}$ $(n = 1, 2, \cdots)$ contains a subsequence which converges weakly to some $\bar{x} \in L(\Omega)$. Since any $x \in L(\Omega)$ can be represented as a difference of two positive elements: $x = x_+ - x_-$, we have only to discuss the case $x \geqq 0$, and our theorem is reduced to the following

THEOREM 10. *Let $\{x_n(t)\}$ $(n = 1, 2, \cdots)$ be a sequence of non-negative measurable functions from $L(\Omega)$. (Ω may not be a sum of a countable infinite number of subsets of finite measure.) If there exists a function $x_0(t) \in L(\Omega)$ such that $x_n(t) \leqq x_0(t)$ almost everywhere on Ω, then the sequence $\{x_n(t)\}$ $(n = 1, 2, \cdots)$ contains a subsequence which converges weakly to some function $\bar{x}(t) \in L(\Omega)$.*

PROOF: When Ω is the interval $(0, 1)$ or $(-\infty, +\infty)$, this theorem is well-known, and the general case can be reduced to this case. To show this, consider the family $\mathbf{K}$ of all the sets $E_{n,\alpha} = E_t[x_n(t) > \alpha]$, where α is any positive number and $n = 0, 1, 2, \cdots$. Each $E_{n,\alpha}$ is of finite measure since we have $\alpha m(E_{n,\alpha}) \leqq \int_\Omega x_n(t)\, dt < \infty$ for any $\alpha > 0$ and $n = 0, 1, 2, \cdots$. Moreover, since we have $E_{n,\alpha} = \lim_{k\to\infty} E_{n,r_k}$, where α is any positive number and $\{r_k\}$ $(k = 1, 2, \cdots)$ is a monotone decreasing sequence of positive rational numbers which tends to α, $\mathbf{K}$ is separable with respect to the metric $d(E_1, E_2) = m(E_1 + E_2 - E_1E_2)$. Hence, by a well-known argument, $\mathbf{K}$ may be represented isometrically and lattice-isomorphically by a family of measurable sets in the interval $(-\infty, +\infty)$.

Thus the characteristic functions of the sets of $\mathbf{K}$ and consequently the functions $\{x_n(t)\}$ $(n = 0, 1, 2, \cdots)$ may be considered as the non-negative measurable functions of $L(\Omega_0)$, where Ω_0 is the infinite interval $(-\infty, +\infty)$. It is clear that we have $x_n(t) \leqq x_0(t)$ $(n = 1, 2, \cdots)$ almost everywhere on Ω_0. Hence by a well-known result, the sequence $\{x_n(t)\}$ $(n = 1, 2, \cdots)$ contains a subsequence $\{x_{n_\nu}(t)\}$ $(\nu = 1, 2, \cdots)$ which converges weakly to a function $\bar{x}(t) \in L(\Omega_0)$. Since $\bar{x}(t)$ belongs to the smallest closed linear manifold which contains $\{x_n(t)\}$ $(n = 1, 2, \cdots)$, $\bar{x}(t)$ may also be considered as to belong to $L(\Omega)$, and it is again clear that the sequence $\{x_{n_\nu}(t)\}$ $(\nu = 1, 2, \cdots)$ converges weakly to $\bar{x}(t)$ as a sequence of elements of (AL).

Thus the proof of Theorem 10 and thereby the proof of Theorem 9 is completed.

REMARK 1. The proof given above is based on the representation theorem (Theorem 7), and in proving Theorem 7 we have made use of the transfinite induction (see Theorem 2). In order to avoid such a transfinite method, we shall prove the following Theorem 11. As will be shown later, Theorem 11 will give us another proof of Theorem 9. Theorem 11 is interesting in itself and will perhaps become a useful tool in the allied problems.

THEOREM 11. *Let (AL) be an abstract (L)-space (satisfying the conditions (I)–(IX)) which is not necessarily separable, and let $\{x_n\}$ $(n = 1, 2, \cdots)$ be an arbitrary sequence of points from (AL). Then there exists a separable closed linear subspace $(AL)'$ of (AL) which contains $\{x_n\}$ $(n = 1, 2, \cdots)$ and which is also closed in the sense of lattice (i.e., $y_1, y_2 \in (AL)'$ implies $y_1 \vee y_2$, $y_1 \wedge y_2 \in (AL)'$).*

Proof: Let Y be the set of all $y \in (AL)$ which is obtained from $\{x_n\}$ ($n = 1, 2, \cdots$) by operating a finite number of times the following operations:

$$\text{taking a sum: } y = y_1 + y_2, \tag{6.1}$$

$$\text{multiplication by a rational number: } y = \lambda y_1, \tag{6.2}$$

$$\text{taking a maximum: } y = y_1 \vee y_2, \tag{6.3}$$

$$\text{taking a minimum: } y = y_1 \wedge y_2. \tag{6.4}$$

Y is clearly a countable set, and $y_1, y_2 \in Y$ implies $y_1 + y_2$, λy_1, $y_1 \vee y_2$, and $y_1 \wedge y_2 \in Y$, where λ is a rational number. Let now $(AL)'$ be a closed cover of Y (in the strong topology of (AL)). We shall prove that this $(AL)'$ is the required one. Since it is clear that $(AL)'$ is separable and contains $\{x_n\}$ ($n = 1, 2, \cdots$), we have only to prove that $(AL)'$ is linear and is closed in the sense of lattice.

For this purpose, let $y_1, y_2 \in (AL)'$ and let λ be an arbitrary real number. Then there exist two sequences of points $\{y_{in}\}$ ($n = 1, 2, \cdots; i = 1, 2$) from Y and a sequence of rational numbers $\{\lambda_n\}$ ($n = 1, 2, \cdots$) such that $y_{in} \to y_i$ (strongly, $i = 1, 2$) and $\lambda_n \to \lambda$. Consequently, $y_{1n} + y_{2n} \to y_1 + y_2$ (strongly), $\lambda_n y_{1n} \to \lambda y_1$ (strongly), $y_{1n} \vee y_{2n} \to y_1 \vee y_2$ (strongly) and $y_{1n} \wedge y_{2n} \to y_1 \wedge y_2$ (strongly). Since $y_{1n} + y_{2n}$, $\lambda_n y_{1n}$, $y_{1n} \vee y_{2n}$ and $y_{1n} \wedge y_{2n}$ belong to Y for $n = 1, 2, \cdots$, $y_1 + y_2$, λy_1, $y_1 \vee y_2$ and $y_1 \wedge y_2$ must belong to $(AL)'$ which completes the proof of Theorem 11.

Proof of Theorem 9. By Theorem 11, there exists a separable closed linear subspace $(AL)'$ of (AL) which contains $\{x_n\}$ ($n = 0, 1, 2, \cdots$) and which is also closed in the sense of lattice. It is clear that $(AL)'$ itself is also an abstract (L)-space. Hence $(AL)'$ can be embedded isometrically and lattice-isomorphically *into* the concrete (L)-space (L) by Theorem 8. Thus $\{x_n\}$ ($n = 0, 1, 2, \cdots$) may be considered as a sequence of measurable functions $\{x_n(t)\}$ ($n = 0, 1, 2, \cdots$) of (L). If we now consider the case $x \geqq 0$, then $0 \leqq x_n(t) \leqq x_0(t)$ almost everywhere for $n = 1, 2, \cdots$. Hence, as we have observed above, there exists a subsequence $\{x_{n_\nu}\}$ ($\nu = 1, 2, \cdots$) of $\{x_n\}$ ($n = 1, 2, \cdots$) which converges weakly (as a sequence of points of (L)) to a point $\bar{x} \in (L)$. Since $\bar{x}$ belongs to a closed linear manifold which is spanned by $\{x_n\}$ ($n = 1, 2, \cdots$), $\bar{x}$ may also be considered as to belong to $(AL)'$, and it is clear that the sequence $\{x_{n_\nu}\}$ ($\nu = 1, 2, \cdots$) converges weakly to $\bar{x}$ as a sequence of points of $(AL)'$. Thus we have proved that there exists a subsequence $\{x_{n_\nu}\}$ ($\nu = 1, 2, \cdots$) of $\{x_n\}$ ($n = 1, 2, \cdots$) which converges weakly to $\bar{x} \in (AL)'$ as a sequence of points of $(AL)'$, and it is again clear that the sequence $\{x_{n_\nu}\}$ ($\nu = 1, 2, \cdots$) converges weakly to $\bar{x}$ as a sequence of points of (AL).

The rest of the proof may now be carried out exactly as in the preceding case. We have only to apply the mean ergodic theorem in Banach spaces (K. Yosida [15], S. Kakutani [7] or F. Riesz [10]).

Remark. Since the space $L(\Omega_0)$, where Ω_0 is the interval $(0, 1)$ or $(-\infty, +\infty)$,

is weakly complete, we can prove in the same manner that each concrete (L)-space $L(\Omega)$ is weakly complete. Hence, by Theorem 7,

THEOREM 12. *Every abstract (L)-space is weakly complete.*

For the case of the space of functions of bounded variation, this theorem was obtained by S. Banach and S. Mazur [2].

MATHEMATICAL INSTITUTE, OSAKA IMPERIAL UNIVERSITY

BIBLIOGRAPHY

1. S. Banach: Théorie des opérations linéaires, Warsaw, 1932.
2. S. Banach and S. Mazur: Zur Theorie der linearen Dimensionen, Studia Math., 4 (1933), 100–112.
3. G. Birkhoff: Dependent probabilities and the space (L), Proc. Nat. Acad. U. S. A., 24 (1938), 154–159.
4. H. Freudenthal: Teilweise geordnete Moduln, Proc. Acad. Amsterdam, 39 (1936), 641-651.
5. E. Hopf: Ergodentheorie, Ergebnisse der Math. und ihrer Grenzgebiete, 5, Berlin, 1937.
6. S. Kakutani: Iteration of bounded linear operations in complex Banach spaces, Proc. Imp. Acad. Japan, 14 (1938), 295–300.
7. S. Katutani: Mean ergodic theorem in abstract (L)-spaces, Proc. Imp. Acad. Japan, 15 (1939), 121–123.
8. L. Kantorovitch: Lineare halbgeordnete Räume, Recueil math., 44 (1937), 121–165.
9. O. Nikodym: Sur une généralisation des intégrales de M. Radon, Fund. Math., 15 (1930), 131–179.
10. F. Riesz: Some mean ergodic theorems, Journal of the London Math. Soc., 13 (1938), 274–278.
11. S. Saks: Theory of the integral, Warsaw, 1937.
12. M. H. Stone: Application of the theory of Boolean rings to general topology, Trans. Amer. Math. Soc., 41 (1937), 376–481.
13. H. Wallman: Lattices and topological spaces, Annals of Math., 39 (1938), 112-126.
14. F. Wecken: Unitärinvarianten selbstadjungierter Operatoren, Math. Ann., 116 (1938), 422-455.
15. K. Yosida: Mean ergodic theorem in Banach spaces, Proc. Imp. Acad. Japan, 14 (1938), 292-294.

Annals of Mathematics
Vol. 42, No. 2, April, 1941

TWO ESTIMATES CONNECTED WITH THE (α, β)-HYPOTHESIS[1]

By Peter Scherk

(Received April 22, 1940)

Small italics denote integers. The "number function" $D(x)$ of a set D of positive integers d is the number of the $d \leqq x$.

Let A, B, C be sets of positive integers a, b, c respectively and $A(x)$, $B(x)$, $C(x)$ their respective number functions. Let C be the set of all the numbers of the form a, b or $a + b$.

The so-called (α, β)-hypothesis reads: Let α and β be two positive real numbers with $\alpha + \beta < 1$. Let $A(x) \geqq \alpha x$ and $B(x) \geqq \beta x$ for $x = 1, 2, \cdots, n$. Then $C(n) \geqq (\alpha + \beta)n$.

This hypothesis has been treated among others by Landau, Besicovitch, and Schur.[2] The proofs of their theorems have the principle in common that the number function $C(n)$ is estimated more or less explicitly through integer expressions in the number functions of the two sets A and B; and it is only afterwards that estimates of these number functions are used. An analysis of Besicovitch's paper shows that Landau's and Schur's theorems can be derived from the estimate contained in his proof.[3]

The problem of constructing such integral estimates of $C(n)$ through expressions in the number functions of A and B seems interesting for two reasons: On the one hand one can hope to reach a deeper understanding of the (α, β)-hypothesis. On the other hand this problem seems more natural, since it involves no hypothesis about the sets A and B.

Until now two estimates of this kind are known besides that of Besicovitch.[3] These three estimates are sharp. In each of them the number of terms can be made arbitrarily large by increasing n.

In this paper two new estimates are presented each containing only a restricted number of terms. Like the earlier three inequalities they are valid for numbers

---

[1] This paper was presented to the American Mathematical Society on February 24, 1940.

[2] E. Landau: Die Goldbachsche Vermutung und der Schnirelmannsche Satz, Nachr. Ges. Wiss. Goettingen, 1930, pp. 255–276.

A. S. Besicovitch: On the density of the sum of two sequences of integers, Journal London Math. Soc. *10* (1935), pp. 246–248.

I. Schur: Ueber den Begriff der Dichte in der additiven Zahlentheorie, Sitz. Ber. Preuss. Akad. Wiss., 1936, pp. 269–297.

Also: E. Landau: Ueber einige neuere Fortschritte der additiven Zahlentheorie, Cambridge Tracts *35* (1937).

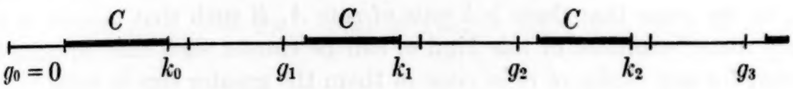
[3] P. Scherk: Bemerkungen zu einer Note von Besicovitch, Journal London Math. Soc. *14* (1939), pp. 185–192.

not lying in C and interconnected by Khintchine's "Umkehrformel" (Inversion formula).[4]

We decompose the positive numbers into intervals, such that all the integers in intervals of one kind belong to C, those in the others not: Let

$$g_0 = 0 \leqq k_0 < g_1 < k_1 < g_2 < k_2 < \cdots ;$$

$$z \subset C \quad \text{for} \quad g_i < z \leqq k_i \, ; \quad z \not\subset C \quad \text{for} \quad k_i < z \leqq g_{i+1} \qquad [i = 0, 1, 2, \cdots].$$



We abbreviate $\Gamma(z) = A(z) + B(z)$.

If $k_0 \leqq n \leqq g_1$, then $C(n) \geqq \Gamma(n)$. The proof of this statement is simple.[5] Two facts follow from it: 1) The (α, β)-hypothesis is true for $k_0 < n \leqq g_1$ [and therefore up to k_1]. 2) If we suppose $\alpha + \beta \geqq 1$ instead of $\alpha + \beta < 1$, then C contains all the positive integers up to n.

If we go from the first segment $k_0 \leqq n \leqq g_1$ to any higher one $k_i \leqq n \leqq g_{i+1}$ $[i \geqq 1]$, then there exists no longer a general inequality of the form $C(n) \geqq \Gamma(n) - f(i)$, where $f(i)$ depends upon i but not on A and B.[6]

The following theorem, however, can be proved: If $k_i \leqq n \leqq g_{i+1}$, and if $C(n) < \Gamma(n)$, then there is a number $l + 1 \leqq k_i$ belonging to both A and B such that

$$C(n) \geqq \Gamma(l) + \Gamma(n - l - 1) - (i - 2).$$

This estimate contains the (α, β)-theorem for the interval $k_1 < n \leqq g_2$. For such an n satisfies at least one of the following two inequalities

$$C(n) \geqq \Gamma(n) \quad \text{and} \quad C(n) \geqq \Gamma(l) + \Gamma(n - l - 1) + 1.$$

The second theorem is connected with the analogous decomposition of the positive integers according to whether or not they belong to B. It can be formulated as follows: Let

$$n + 1 \not\subset C, \qquad C(n) < A(n) + B(n). \tag{1}$$

Then there is a number $m < n$ with

$$m + 1 \not\subset C, \qquad n - m \subset A,$$

---

[4] A. Khintchine: Zur additiven Zahlentheorie, Matem. Sbornik *39*, 3 (1932), pp. 27–34. One of the inequalities being symmetrical, only one other estimate is related to it.

[5] See the beginning of the proof of the first theorem. This fact was already used by Schnirelmann. Schnirelmann: Ueber additive Eigenschaften der Zahlen, Math. Annalen *107* (1933), pp. 649–690.

[6] One can readily find examples which show that $\Gamma(n) - C(n)$ can be made arbitrarily great while i remains fixed: Let

$$1 \leqq r < v, \qquad A = B = \{1, 2, \cdots, r, 2v + 1, 2v + 2, \cdots, 3v\}.$$

Hence $C = \{1, 2, \cdots, 2r, 2v + 1, 2v + 2, \cdots, 3v + r, 4v + 2, 4v + 3, \cdots, 6v\}$ $k_1 = 3v + r$ and $\Gamma(k_1) - C(k_1) = 2(v + r) - (2r + v + r) = v - r$.

such that

$$(2)\qquad \begin{aligned}&(C(n) - C(n-m)) + (C(n) - C(m+1))\\ &\qquad \geqq A(m+1) + A(n-m-1) + B(n) - (i-1).\end{aligned}$$

Here $i + 1$ equals the number of numbers $r \leqq n$ with $r + 1 \not\subset B$ and either $r \subset B$ or $r = 0$.

Both theorems can be slightly refined, but they are already sharp in the given form, in the sense that there is a pair of sets A, B such that in each of the infinitely many estimates of this kind m can be chosen such that equality holds, but that for any choice of m in none of them the greater sign is valid.[7]

Throughout this paper we abbreviate $D(x, y) = D(y) - D(x)$ for any number function $D(x)$[8] and $\Gamma(x, y) = \Gamma(y) - \Gamma(x)$. Thus (2) can be written

$$(3)\qquad \begin{aligned}&C(n-m, n) + C(m+1, n)\\ &\qquad \geqq A(m+1) + A(n-m-1) + B(n) - (i-1).\end{aligned}$$

Lemma:[9] Let $k + 1 \not\subset C$, $0 \leqq g < h \leqq k$. Then

$$(4)\qquad h - g \geqq A(g, h) + B(k - h, k - g).$$

Proof: Let $a \subset A$, $b \subset B$. From $a = k + 1 - b$ would follow $k + 1 = a + b \subset C$. Therefore, the numbers a with $g < a \leqq h$ and the numbers $k + 1 - b$ with $g < k + 1 - b \leqq h$ are different from each other. Their number

$$A(g, h) + B(k - h, k - g)$$

is not greater than the number $h - g$ of all the numbers z with $g < z \leqq h$.

Upon adding the formula symmetrical to it to (4), we obtain

$$(5)\qquad 2(h - g) \geqq \Gamma(g, h) + \Gamma(k - h, k - g).$$

Proof of the first theorem:

The k_i and g_i have the same meaning as in the introduction. If $k_0 > 0$, then

$$C(k_0) = k_0 \geqq \Gamma(k_0)$$

according to (4) with $g = 0$, $h = k = k_0$. If $k_0 = 0$, this estimate is trivial. If

$$(6)\qquad C(x) \geqq \Gamma(x)$$

and if $x + 1 \not\subset C$, then (6) also holds for $x + 1$ instead of x.

We may suppose (6) is not valid for all the $x \not\subset C$. Then there is a $j_1 \geqq 1$ such that (6) holds for all the $x \leqq g_{j_1}$ with $x \not\subset C$ but not for $x = k_{j_1}$. We shall show that there exists an inequality of the desired kind for all the $x \geqq k_{j_1}$ belonging to an interval $k_i \leqq x \leqq g_{i+1}$.

---

[7] Cf. the example given in 3.

[8] When $x < y$, $D(x, y)$ obviously indicates the number of the $d \subset D$ with $x < d \leqq y$.

[9] This lemma is a well known special case of Khintchine's "Umkehrformel"; cf. 2, 3, 4.

There is a number l_1 with $g_{j_1} \leqq l_1 < k_{j_1}$ such that

$$\Gamma(l_1) \leqq C(l_1) < C(l_1 + 1) < \Gamma(l_1 + 1). \tag{7}$$

Thus $l_1 + 1$ must lie in the intersection $A \cdot B$ of A and B. The formula (5) with

$$k = k_{j_1}, \qquad g = 0, \qquad h = k_{j_1} - l_1$$

gives

$$2C(l_1, k_{j_1}) = 2(k_{j_1} - l_1) \geqq \Gamma(k_{j_1} - l_1) + \Gamma(l_1, k_{j_1})$$

hence, because of (7)

$$\begin{aligned} 2C(k_{j_1}) = 2C(l_1) + 2C(l_1, k_{j_1}) &\geqq 2\Gamma(l_1) + \Gamma(k_{j_1} - l_1) + \Gamma(l_1, k_{j_1}) \\ &\geqq \Gamma(l_1) + \Gamma(k_{j_1}) + \Gamma(k_{j_1} - l_1 - 1). \end{aligned}$$

Since $C(k_{j_1}) \leqq \Gamma(k_{j_1}) - 1$ we obtain

$$C(k_{j_1}) \geqq \Gamma(l_1) + \Gamma(k_{j_1} - l_1 - 1) + 1.$$

The numbers $k_{j_1}, k_{j_2}, \cdots, k_{j_i}$ and $l_1, l_2, \cdots, l_i$ may already have been constructed $[i \geqq 1]$ such that $l_i + 1 \subset A \cdot B$ and

$$C(x) \geqq \Gamma(l_i) + \Gamma(x - l_i - 1) - (i - 2) \tag{8}$$

for $x = k_{j_i}$. If (8) holds for x and if $x + 1 \not\subset C$, then (8) also holds for $x + 1$ instead of x on account of $l_i + 1 \subset A \cdot B$.

If (8) is not valid for all the $x \not\subset C$ with $x > k_{j_i}$, then there exists a number $j_{i+1} > j_i$ such that (8) holds for all the $x \not\subset C$ with $k_{j_i} < x \leqq g_{j_{i+1}}$ but not for $x = k_{j_{i+1}}$; and there is a number m with $g_{j_{i+1}} \leqq m < k_{j_{i+1}}$ such that

$$\begin{cases} \Gamma(l_i) + \Gamma(m - l_i - 1) - (i - 2) \leqq C(m) < C(m + 1) \\ \qquad\qquad < \Gamma(l_i) + \Gamma(m - l_i) - (i - 2). \end{cases} \tag{9}$$

Therefore

$$l_{i+1} + 1 = m - l_i \subset A \cdot B.$$

From (5) with

$$k = k_{j_{i+1}}, \qquad g = l_{i+1} + 1, \qquad h = k_{j_{i+1}} - l_i$$

we have

$$\begin{aligned} 2C(m, k_{j_{i+1}}) = 2(k_{j_{i+1}} - m) &= 2((k_{j_{i+1}} - l_i) - (l_{i+1} + 1)) \\ &\geqq \Gamma(l_{i+1} + 1, k_{j_{i+1}} - l_i) + \Gamma(l_i, k_{j_{i+1}} - l_{i+1} - 1) \\ &\geqq \Gamma(l_{i+1}, k_{j_{i+1}} - l_i - 1) + \Gamma(l_i, k_{j_{i+1}} - l_{i+1} - 1) - 2. \end{aligned}$$

On account of (9) we obtain

$$\begin{aligned} 2C(k_{j_{i+1}}) = 2C(m) + 2C(m, k_{j_{i+1}}) &\geqq 2(\Gamma(l_i) + \Gamma(l_{i+1}) - (i - 2)) + \\ &\quad + \Gamma(l_{i+1}, k_{j_{i+1}} - l_i - 1) + \Gamma(l_i, k_{j_{i+1}} - l_{i+1} - 1) - 2 \\ &= \Gamma(l_i) + \Gamma(l_{i+1}) + \Gamma(k_{j_{i+1}} - l_i - 1) + \Gamma(k_{j_{i+1}} - l_{i+1} - 1) - 2(i - 1); \end{aligned}$$

and since (8) should not be true for $x = k_{j_{i+1}}$

$$C(k_{j_{i+1}}) \geqq \Gamma(l_{i+1}) + \Gamma(k_{j_{i+1}} - l_{i+1} - 1) - (i - 1).$$

That accomplishes the induction, and $j_i \geqq i$ gives our theorem.[10]

Proof of the second theorem:

In the following a, b, c denote numbers of A, B, C respectively.

The number $n + 1$ satisfies the condition (1). Let

$$n = s_0 \geqq r_0 > s_1 > r_1 > s_2 > \cdots > r_{i-1} > s_i > r_i \geqq s_{i+1} = 0,$$

$$x \subset B \quad \text{for} \quad s_{j+1} < x \leqq r_j\,; \quad x \not\subset B \quad \text{for} \quad r_j < x \leqq s_j \quad [j = 0, 1, 2, \cdots, i].$$

From $a \leqq n - r_0$ follows $r_0 < c = a + r_0 \leqq n$; therefore

$$C(r_0\,, n) \geqq A(n - r_0) = A(n - r_0) + B(r_0\,, n).$$

Let

$$C(x, n) \geqq A(n - x) + B(x, n) \tag{10}$$

hold for $x = r_\lambda$. Then (10) is true for all the x with $s_{\lambda+1} \leqq x \leqq r_\lambda$. For (4) with

$$k = n, \qquad g = n - r_\lambda\,, \qquad h = n - x$$

---

[10] If $k_{j_1} < x < k_{j_2}$, $x \not\subset C$ then

$$C(x) \geqq \Gamma(g_{j_1}) + \Gamma(x - g_{j_1} - 1) + 1. \tag{1'}$$

Since in view of (8)

$$C(x) \geqq \Gamma(l_1) + \Gamma(x - l_1 - 1) + 1, \tag{2'}$$

it is sufficient to show that the right term of (2′) is not smaller than that of (1′). Thus (1′) follows from

$$2\Gamma(g_{j_1}\,, l_1) \geqq \Gamma(g_{j_1}\,, l_1) + \Gamma(x - l_1 - 1, x - g_{j_1} - 1). \tag{3'}$$

Since, on account of the lemma, the right term of (3′) is not greater than $2C(g_{j_1}\,, l_1)$, we have only to show

$$\Gamma(g_{j_1}\,, l_1) \geqq C(g_{j_1}\,, l_1).$$

But this is evident; for, on account of (7), $C(l_1) = \Gamma(l_1)$ and $C(g_{j_1}) \geqq \Gamma(g_{j_1})$.

The following example shows that there is no estimate analogous to (2′), from which the (α, β)-hypothesis would result for the interval $k_{j_2} < x \leqq g_{j_2+1}$:

$$A = \{1, 2, \cdots, t - 1, 2t, 2t + 1, \cdots, 3t - 1, 4t, 4t + 1, \cdots, 5t - 1\}$$
$$B = \{1, 2, \cdots, t - 1, 2t, 2t + 1, \cdots, 3t - 1, 4t + 1, 4t + 2, \cdots, 5t - 1\} \qquad [t \geqq 2].$$

There has not been found any example showing the same for (1′).

gives

$$\begin{aligned} C(x, n) = C(x, r_\lambda) + C(r_\lambda, n) &\geqq r_\lambda - x + A(n - r_\lambda) + B(r_\lambda, n) \\ &\geqq A(n - r_\lambda, n - x) + B(x, r_\lambda) + A(n - r_\lambda) + B(r_\lambda, n) \\ &= A(n - x) + B(x, n). \end{aligned}$$

On account of (1) there exists, therefore, a number j_1 with $1 \leqq j_1 \leqq i$ such that (10) is valid for $x = r_\lambda$ $[\lambda = 0, 1, \cdots, j_1 - 1]$ but not for $x = r_{j_1}$. Then (10) holds especially for $x = s_{j_1}$. Let m_1 be the greatest number x with $r_{j_1} \leqq x < s_{j_1}$ for which (10) is not true; thus

$$\begin{aligned} A(n - m_1) + B(m_1, n) > C(m_1, n) \geqq C(m_1 + 1, n) \\ \geqq A(n - m_1 - 1) + B(m_1 + 1, n). \end{aligned} \tag{11}$$

Since

$$B(m_1, n) = B(m_1 + 1, n) = B(r_{j_1}, n)$$

we have

$$m_1 + 1 \not\subset C, \qquad n - m_1 \subset A.$$ [11]

Furthermore

$$C(r_{j_1}, m_1) \geqq A(m_1 - r_{j_1}), \quad C(n - m_1 + r_{j_1}, n) \geqq A(n - m_1, n - r_{j_1}). \tag{12}$$

For if $r_{j_1} \neq 0$, then $r_{j_1} \subset B$;

$$a \leqq m_1 - r_{j_1} \quad \text{gives} \quad r_{j_1} < c = a + r_{j_1} \leqq m_1$$

and $n - m_1 < a \leqq n - r_{j_1}$ gives $n - m_1 + r_{j_1} < c = a + r_{j_1} \leqq n.$

Thus we obtain:

$$\begin{aligned} C(n - m_1 + r_{j_1}, n) + C(m_1 + 1, n) &\geqq A(n - m_1, n - r_{j_1}) + \\ &\quad + A(n - m_1 - 1) + B(m_1 + 1, n) \qquad \text{[see (11) and (12)]} \\ &\geqq A(n - r_{j_1}) - 1 + B(r_{j_1}, n) \\ &\geqq C(r_{j_1}, n) \qquad \text{[(10) should not hold for } x = r_{j_1}] \\ &= C(r_{j_1}, m_1) + C(m_1, n) \\ &\geqq A(m_1 - r_{j_1}) + A(n - m_1 - 1) + B(r_{j_1}, n) \qquad \text{[see (11) and (12)].} \end{aligned}$$

Let m_κ and r_{j_κ} be already defined,

$$0 < \kappa \leqq j_\kappa < i, \qquad r_{j_\kappa} \leqq m_\kappa < n,$$

$$m_\kappa + 1 \not\subset C, \qquad n - m_\kappa \subset A,$$

[11] By iterating the step leading to m_1 a descending sequence can be constructed which is essentially equivalent to Besicovitch's ascending one.

and let

$$C(n - m_\kappa + x, n) + C(m_\kappa + 1, n) \geqq A(m_\kappa - x) + A(n - m_\kappa - 1) + B(x, n) - (\kappa - 1) \tag{13}$$

hold for $x = r_{j_\kappa}$.

If (13) is true for $x = r_\lambda$, then it holds for all the x with $s_{\lambda+1} \leqq x < r_\lambda$. Since on account of $n - m_\kappa \subset A$ and $z \subset B$ for $s_{\lambda+1} < z \leqq r_\lambda$ every number $n - m_\kappa + z$ belongs to C, we have

$$C(n - m_\kappa + x, n - m_\kappa + r_\lambda) = r_\lambda - x;$$

from (13) with r_λ instead of x and (4) with

$$k = m_\kappa, \qquad g = m_\kappa - r_\lambda, \qquad h = m_\kappa - x$$

we hence derive

$$\begin{aligned} C(n - m_\kappa + x, n) &+ C(m_\kappa + 1, n) \\ &= r_\lambda - x + C(n - m_\kappa + r_\lambda, n) + C(m_\kappa + 1, n) \\ &\geqq A(m_\kappa - r_\lambda, m_\kappa - x) + B(x, r_\lambda) + A(m_\kappa - r_\lambda) + \\ &\quad + A(n - m_\kappa - 1) + B(r_\lambda, n) - (\kappa - 1) \\ &= A(m_\kappa - x) + A(n - m_\kappa - 1) + B(x, n) - (\kappa - 1). \end{aligned}$$

If we can choose in particular $\lambda = i$, (13) holds for $x = s_{i+1} = 0$ and, on account of $\kappa \leqq i$ and $A(m_\kappa) = A(m_\kappa + 1)$, we arrive at the assertion (3) with $m = m_\kappa$.

We suppose now that (13) is not true for $x = 0$. Then there is a number $j_{\kappa+1}$ with $j_\kappa < j_{\kappa+1} \leqq i$ such that (13) holds for $x = r_\lambda$ $[\lambda = j_\kappa, j_\kappa + 1, \cdots, j_{\kappa+1} - 1]$ but not for $x = r_{j_{\kappa+1}}$. Thus we have

$$A(m_\kappa - r_{j_{\kappa+1}}) + A(n - m_\kappa - 1) + B(r_{j_{\kappa+1}}, n) - \kappa \geqq C(n - m_\kappa + r_{j_{\kappa+1}}, n) + C(m_\kappa + 1, n). \tag{14}$$

According to the above-mentioned (13) holds for $x = s_{j_{\kappa+1}}$. Let l be the greatest x with $r_{j_{\kappa+1}} \leqq x < s_{j_{\kappa+1}}$, for which (13) does not hold. Then we have

$$\left\{\begin{aligned} &A(m_\kappa - l) + A(n - m_\kappa - 1) + B(l, n) - \kappa \\ &\qquad \geqq C(n - m_\kappa + l, n) + C(m_\kappa + 1, n) \\ &\qquad \geqq C(n - m_\kappa + l + 1, n) + C(m_\kappa + 1, n) \\ &\qquad \geqq A(m_\kappa - l - 1) + A(n - m_\kappa - 1) + B(l + 1, n) - (\kappa - 1), \end{aligned}\right. \tag{15}$$

$$B(l, n) = B(l + 1, n) = B(r_{j_{\kappa+1}}, n).$$

We put

$$m_{\kappa+1} = n - m_\kappa + l,$$

thus

$$r_{j_{\kappa+1}} \leqq l < (n - m_\kappa) + l = m_{\kappa+1} = n - (m_\kappa - l) < n,$$

and (15) can be written

$$(16)\quad \begin{cases} A(n - m_{\kappa+1}) + A(n - m_\kappa - 1) + B(r_{j_{\kappa+1}}, n) - \kappa \\ \qquad \geqq C(m_{\kappa+1}, n) + C(m_\kappa + 1, n) \\ \qquad \geqq C(m_{\kappa+1} + 1, n) + C(m_\kappa + 1, n) \\ \qquad \geqq A(n - m_{\kappa+1} - 1) + A(n - m_\kappa - 1) + B(r_{j_{\kappa+1}}, n) - (\kappa - 1). \end{cases}$$

From (16)

$$m_{\kappa+1} + 1 \not\subset C, \qquad n - m_{\kappa+1} \subset A.$$

Further

$$n - m_{\kappa+1} + r_{j_{\kappa+1}} = m_\kappa - l + r_{j_{\kappa+1}} < m_\kappa .$$

Since $r_{j_{\kappa+1}} = 0$ or $r_{j_{\kappa+1}} \subset B$ we have in analogy to (12)

$$(17)\quad \begin{cases} C(n - m_{\kappa+1} + r_{j_{\kappa+1}}, m_\kappa) \geqq A(n - m_{\kappa+1}, m_\kappa - r_{j_{\kappa+1}}) \\ C(n - m_\kappa + r_{j_{\kappa+1}}, m_{\kappa+1}) \geqq A(n - m_\kappa, m_{\kappa+1} - r_{j_{\kappa+1}}). \end{cases}$$

Therefore

$$\begin{aligned} &C(m_{\kappa+1} + 1, n) + C(n - m_{\kappa+1} + r_{j_{\kappa+1}}, n) \\ &\quad = C(m_{\kappa+1} + 1, n) + C(m_\kappa + 1, n) + C(n - m_{\kappa+1} + r_{j_{\kappa+1}}, m_\kappa) \\ &\quad \geqq A(n - m_{\kappa+1} - 1) + A(n - m_\kappa - 1) + B(r_{j_{\kappa+1}}, n) - \\ &\quad - (\kappa - 1) + A(n - m_{\kappa+1}, m_\kappa - r_{j_{\kappa+1}}) \quad \text{[see (16) and (17)]} \\ &\quad \geqq A(m_\kappa - r_{j_{\kappa+1}}) + A(n - m_\kappa - 1) + B(r_{j_{\kappa+1}}, n) - \kappa \\ &\quad \geqq C(n - m_\kappa + r_{j_{\kappa+1}}, n) + C(m_\kappa + 1, n) \\ &\quad = C(n - m_\kappa + r_{j_{\kappa+1}}, m_{\kappa+1}) + C(m_{\kappa+1} + 1, n) + C(m_\kappa + 1, n) \\ &\quad \geqq A(n - m_\kappa, m_{\kappa+1} - r_{j_{\kappa+1}}) + A(n - m_{\kappa+1} - 1) + A(n - m_\kappa - 1) + \\ &\quad + B(r_{j_{\kappa+1}}, n) - (\kappa - 1) \\ &\quad \geqq A(m_{\kappa+1} - r_{j_{\kappa+1}}) + A(n - m_{\kappa+1} - 1) + B(r_{j_{\kappa+1}}, n) - \kappa. \end{aligned}$$

That finishes the induction.[12]

YALE UNIVERSITY.

[12] In analogy to footnote 10 one can prove: Let $s_{j_2} < x < s_{j_1}$, $x \subset B$; then $C(n - s_{j_1} + x, n) + C(s_{j_1}, n) \geqq A(n - s_{j_1}) + A(s_{j_1} - x) + B(x, n) + 1$. From (11) follows

$$(4') \qquad C(m_1, n) = A(n - m_1 - 1) + B(m_1, n) = A(n - m_1) + B(s_{j_1}, n) - 1.$$

From (13) with $\kappa = 1$ and (4′) follows

$$(5') \qquad C(n - m_1 + x, n) \geqq A(m_1 - x) + B(x, s_{j_1}).$$

Furthermore according to (10) with s_{j_1} instead of x:

$$(6') \qquad C(s_{j_1}, n) \geqq A(n - s_{j_1}) + B(s_{j_1}, n).$$

(4′) and (6′) together give

$$(7') \qquad C(m_1, s_{j_1}) \leqq A(n - s_{j_1}, n - m_1) - 1.$$

Finally

$$(8') \qquad \begin{cases} C(n - s_{j_1} + x, n - m_1 + x) \geqq A(n - s_{j_1}, n - m_1) & [\text{for } x \subset B] \\ \qquad \geqq C(m_1, s_{j_1}) + 1 & [\text{see } (7')] \\ \qquad \geqq A(m_1 - x, s_{j_1} - x) + 1 & [x \subset B]. \end{cases}$$

By adding (5′), (6′), and (8′) we obtain the assertion.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

SUR LE THÉORÈME DE LEBESGUE-NIKODYM

PAR JEAN DIEUDONNÉ

(Received May 16, 1940)

INTRODUCTION. Un des théorèmes les plus importants de la théorie de l'intégrale de Lebesgue est le suivant: si E désigne l'ensemble des fonctions réelles continues $x(t)$ dans l'intervalle $0 \leq t \leq 1$, toute fonctionelle linéaire positive $L(x)$ définie dans E peut se mettre, d'une manière et d'une seule, sous la forme

$$L(x) = \int_0^1 y(t)x(t)\,dt + S(x) \tag{1}$$

où $y(t)$ est une fonction mesurable positive, bien définie dans l'intervalle $[0, 1]$, à l'exception des points d'un ensemble de mesure nulle, et $S(x)$ une fonctionnelle linéaire positive *singulière*, c'est-à-dire jouissant de la propriété suivante: il existe un ensemble de mesure nulle H, contenu dans l'intervalle $[0, 1]$, tel que, pour toute fonction mesurable positive $z(t)$, nulle en tout point de H, on ait $S(z) = 0$. Ce théorème, généralisé par O. Nikodym à l'intégrale de Radon-Stieltjes[1] (où les fonctions intégrées sont définies dans un ensemble quelconque) est connu dans la littérature sous le nom de *théorème de Lebesgue-Nikodym*.

Dans un important mémoire sur les Opérations linéaires, publié récemment dans ce journal,[2] M. F. Riesz a montré l'extrême généralité de la décomposition qui apparaît dans la formule (1). En prenant comme ensemble E, non plus un ensemble de *fonctions*, mais un ensemble d'éléments de nature quelconque, possédant seulement quelques-unes des propriétés des ensembles de fonctions réelles intervenant dans le théorème de Lebesgue-Nikodym (notamment en ce qui concerne la structure d'*ordre* de cet ensemble, et sa structure de *groupe abélien*), il parvient néanmoins à montrer, par des moyens fort simples, que, si U est une fonction linéaire positive définie sur E, toute autre fonction linéaire positive L sur E peut se mettre, d'une manière et d'une seule, sous la forme $L = V + S$, où V joue le rôle de la fonctionnelle "absolument continue" $\int_0^1 y(t)x(t)\,dt$ du second membre de (1), et S celui de la partie "singulière", de la manière suivante: S est "disjointe" de U, c'est-à-dire que l'on a inf $(U, S) = 0$ (autrement dit, il n'existe pas de fonction linéaire positive, autre que 0, inférieure à la fois à U et S); et V appartient à la plus petite "famille complète" de fonctions linéaires positives contenant U (une telle famille étant un ensemble de fonctions linéaires positives caractérisé par les propriétés suivantes: il contient

[1] Voir par exemple *S. Saks, Theory of the Integral*, New York, G. E. Stechert, 1937.

[2] *F. Riesz, Sur quelques notions fondamentales dans la théorie générale des opérations linéaires*, Annals of Math., *41*, (1940), pp. 174–206.

la somme de deux quelconques de ses éléments, les minorants d'un quelconque de ses éléments, et les bornes supérieures de ses parties majorées).

Toutefois, la théorie de F. Riesz ne permet pas de donner à la partie "absolument continue" de cette décomposition une forme aussi précise que dans la formule (1); cela est dû en premier lieu au fait que le *produit* de deux éléments de l'ensemble E n'est pas défini.

Dans le même temps que je prenais connaissance du mémoire de F. Riesz, N. Bourbaki a bien voulu me communiquer un travail manuscrit[3] sur l'Intégration rédigé en 1939, indépendamment des travaux de F. Riesz, et qui contient, entre autres, une généralisation du théorème de Lebesgue-Nikodym ayant beaucoup de points communs avec les résultats de ce dernier. Chez N. Bourbaki, les éléments de l'ensemble E sont des fonctions comme dans le cas classique, mais sans les restrictions de "mesurabilité" habituelles; en fait, une grande partie de ses résultats n'utilise que les propriétés attribuées par F. Riesz à son ensemble E. En outre, tandis que les structures d'ordre n'interviennent, dans le mémoire de F. Riesz, que par les notions de borne supérieure et borne inférieure, les méthodes de N. Bourbaki reposent sur un usage prépondérant des "limites dans un ensemble ordonné filtrant";[4] le résultat fondamental auquel on aboutit ainsi (et que je reproduis avec l'autorisation de l'auteur) est le suivant: appelons *partition* d'une fonction positive $x \in E$, toute suite finie (x_i) de fonctions positives appartenant à E, et telles que $\sum_i x_i = x$; et *ordonnons* l'ensemble $\mathcal{P}(x)$ des partitions de x, en posant $(x_i) \leq (x'_j)$ si, pour toute fonction x_i de la première partition, il existe une suite finie extraite de la seconde, et qui soit une partition de x_i ; $\mathcal{P}(x)$, ainsi ordonné, est *filtrant* (à droite). Ceci posé, soit $\varphi(u_1, u_2, \cdots, u_p)$ une fonction de p variables réelles, définie dans tout l'espace numérique R^p, *lipschitzienne* (c'est-à-dire telle qu'il existe une constante c satisfaisant à l'identité $|\varphi(u_1, u_2, \cdots, u_p) - \varphi(u'_1, u'_2, \cdots, u'_p)| \leq c\cdot\text{Max}\, |u_k - u'_k|$) et *positivement homogène* (c'est-à-dire que, pour tout $\lambda > 0$, $\varphi(\lambda u_1, \lambda u_2, \cdots, \lambda u_p) = \lambda\cdot\varphi(u_1, u_2, \cdots, u_p)$). Soient d'autre part $I_1, I_2, \cdots, I_p$, p fonctions linéaires "relativement bornées" sur E (différences de deux fonctions linéaires positives); alors, pour tout $x \in E$ et ≥ 0, la formule

$$J(x) = \lim_{\mathcal{P}(x)} \sum_i \varphi(I_1(x_i), I_2(x_i), \cdots, I_p(x_i)) \tag{2}$$

définit une fonction linéaire relativement bornée sur E (la limite étant prise suivant le filtre des sections de l'ensemble ordonné filtrant $\mathcal{P}(x)$); cette fonction se note $\varphi(I_1, I_2, \cdots, I_p)$.[5]

---

[3] Ce manuscrit n'est pas destiné à une publication immédiate, mais constitue un premier projet du fascicule des "*Eléments de Mathématique*" de cet auteur, qui sera consacré à la théorie de l'Intégration. Signalons à ce propos que nous suivons ici la terminologie de ces "Eléments" en ce qui concerne la Théorie des Ensembles et la Topologie générale.

[4] Cette notion de limite n'est autre que celle introduite par E. H. Moore et H. L. Smith ("*A general theory of limits,*" Amer. Journ. of Math., *44* (1922), p. 102).

[5] F. Riesz définit aussi ces "fonctions de fonctions" linéaires, mais d'une manière tout autre et plus détournée.

Ce résultat permet à N. Bourbaki de définir une *topologie d'espace localement convexe* sur l'ensemble F des fonctions linéaires relativement bornées sur E; et, à partir de cette topologie, de donner une nouvelle définition de la "plus petite famille complète" (au sens de F. Riesz) contenant une fonction linéaire positive U. De plus, cette définition topologique donne un moyen de préciser la forme de la partie "absolument continue" de la décomposition de Riesz, en adjoignant de nouveaux éléments à E par une opération de "complétion", et en définissant le produit d'un de ces éléments (qui ne sont plus des fonctions, contrairement aux éléments de E), et d'un élément de E; on montre alors que la partie "absolument continue" de la décomposition de Riesz peut s'écrire $U(yx)$, où y est un des éléments de l'ensemble E "complété" (U étant convenablement "prolongée" à cet ensemble); on arrive ainsi à une généralisation parfaite du théorème de Lebesgue-Nikodym.[6]

Toutefois, cette dernière partie du travail de N. Bourbaki (dont le point capital est la démonstration de l'identité de la partie de F définie par voie topologique, et de la "plus petite famille complète" correspondante) suppose essentiellement que les éléments de E sont des *fonctions*. Il restait à examiner la possibilité d'arriver à des résultats analogues en demeurant dans la voie suivie par F. Riesz, c'est-à-dire sans supposer le caractère fonctionnel des éléments de l'ensemble E; c'est ce que je fais dans ce qui suit. J'y utilise les méthodes topologiques de N. Bourbaki, ainsi que la formule (2) (qui ne suppose pas que les éléments de E sont des fonctions);[7] quant au raccord avec la théorie de F. Riesz, qui reste le point délicat de la démonstration, il se fait en adaptant au cas "abstrait" une idée de J. von Neumann, appliquée par ce dernier à la démonstration du théorème de Lebesgue-Nikodym classique, et qui consiste à passer par l'intermédiaire de l'espace des "fonctions de carré sommable."[8]

1. Nous prenons comme point de départ un ensemble E sur lequel est défini, d'une part une structure d'*ordre*, d'autre part une structure d'*espace vectoriel* par rapport au corps des nombres réels; nous supposons en outre que les conditions suivantes sont remplies:

(I) E est *réticulé* ("lattice"), autrement dit, quels que soient x et y dans E, il existe les éléments $\inf(x, y)$ et $\sup(x, y)$.

(II) La relation $x \leq y$ entraîne $x + z \leq y + z$ quel que soit z.

(III) Les conditions $x \geq 0$, $\lambda \geq 0$ (λ réel) entraînent $\lambda x \geq 0$. On pose $|x| = \sup(x, -x)$; on démontre sans peine les identités $|x + y| \leq |x| + |y|$, $|\lambda x| = |\lambda|\cdot|x|$; si $x^+ = \sup(x, 0)$, $x^- = \sup(-x, 0)$ on a $x = x^+ - x^-$, $|x| = x^+ + x^-$. L'ensemble des éléments ≥ 0 de E sera désigné par E_+ ; il

[6] En ce qui concerne les travaux de N. Bourbaki sur l'Intégration on pourra consulter un article de A. Weil sur la théorie des probabilités (*Revue Scientifique* (Revue rose), 1940.

[7] A l'exception de cette formule, tous les résultats de N. Bourbaki utilisés dans ce travail sont donnés avec leur démonstration.

[8] Voir par exemple, *J. von Neumann, On rings of Operators, III*, Annals of Math., *41* (1940), p. 127-129.

satisfait aux conditions imposées par F. Riesz à son "domaine fondamental" (loc. cit., §I).

Une fonction réelle U sur E est *linéaire* si $U(x + y) = U(x) + U(y)$, et $U(\lambda x) = \lambda U(x)$; elle est *positive* si $U(x) \geq 0$ quel que soit $x \in E_+$; elle est *relativement bornée* si, quel que soit $y \in E_+$, $|U(x)|$ est borné pour l'ensemble des $x \in E$ tels que $|x| \leq y$. Toute fonction relativement bornée est la différence de deux fonctions linéaires positives (N. Bourbaki, loc. cit.).

On désignera par F l'ensemble des fonctions linéaires relativement bornées; si on écrit $U \geq V$ lorsque $U - V \geq 0$, F satisfait aux conditions (I), (II) et (III). D'après la formule (2), on a pour tout $x \in E_+$ et tout $U \in F$, $|U|(x) = \sup \sum_i |U(x_i)|$ pour toutes les *partitions finies* (x_i) de x. On désignera par F_+ l'ensemble des fonctions linéaires positives.

On peut, avec N. Bourbaki, définir sur F une structure d'espace vectoriel *localement convexe*, à l'aide de la famille de *pseudo-normes* $N_x(X) = |X|(x)$ (x fixe dans E_+); F, muni de cette structure, est *complet*. En effet, soit $\mathcal{F}$ un filtre de Cauchy sur F; quel que soit $x \in E_+$, il existe un ensemble $A \in \mathcal{F}$ tel que, pour X et Y quelconques dans A, $|X - Y|(x) \leq \epsilon$, et à fortiori $|X(x) - Y(x)| \leq \epsilon$. Si $A(x)$ désigne l'ensemble des nombres $X(x)$ pour $X \in A$, et $\mathcal{F}(x)$ le filtre formé par les ensembles $A(x)$ sur la droite numérique R, $\mathcal{F}(x)$ est un filtre de Cauchy, qui converge donc vers une limite, qu'on notera $X_0(x)$. Montrons que X_0 est la limite de $\mathcal{F}$ dans F; pour tout $X \in A$, et toute partition finie (x_i) de x, on a

$$\sum_i |X_0(x_i) - X(x_i)| = \sum_i \lim_{\mathcal{F}(x_i)} |Y(x_i) - X(x_i)| \leq \sup_{Y \in A} \sum_i |Y(x_i) - X(x_i)|$$

$$\leq \sup_{Y \in A} |Y - X|(x) \leq \epsilon$$

ce qui montre d'abord que $X_0 - X \in F$, d'où $X_0 \in F$, puis que $|X_0 - X|(x) \leq \epsilon$, ce qui achève la démonstration.

Dans F, tout ensemble *majoré* A possède une borne supérieure (F. Riesz, loc. cit., §II), et toute ensemble minoré une borne inférieure; si en outre A est ordonné *filtrant à droite*, sa borne supérieure est aussi la limite (dans la topologie définie ci-dessus) du *filtre des sections* sur A; on peut d'ailleurs retrouver directement cette propriété en partant du fait que F est complet, et en déduire ensuite le résultat de F. Riesz. Si X_0 est la limite de l'ordonné filtrant A, on a $X_0(x) = \lim_{X \in A} X(x)$, quel que soit $x \in E$.

Remarquons encore que, dans F, F_+ est un ensemble *fermé*, et que les fonctions X^+, X^-, $|X|$ sont *uniformément continues*.

2. Nous supposerons à partir de maintenant que E est muni, en outre des structures précédentes, d'une structure d'*anneau commutatif*, c'est-à-dire que le *produit* xy de deux éléments de E est toujours défini, est associatif, commu-

tatif, distributif par rapport à l'addition et del que $\lambda(xy) = (\lambda x)y$ pour λ réel; enfin, nous faisons l'hypothèse suivante:

(IV) Quels que soient y, z dans E, la relation $x \geq 0$ entraîne

$$\sup(xy, xz) = x \cdot \sup(y, z)$$

Indiquons d'abord quelques conséquences de cette hypothèse. En premier lieu, si $y \leq z$, on a $xy \leq xz$ si $x \geq 0$; en particulier, si $x \geq 0$, $y \geq 0$, $xy \geq 0$; si $x \geq 0$ ou $x \leq 0$, on a $x^2 = (-x)^2 \geq 0$. Si $x \geq 0$ et y est quelconque, $|xy| = x \cdot |y|$; si donc x et y sont quelconques et qu'on pose $x' = |x|$, $y' = |y|$, on aura $|xy| = |x^+y - x^-y| \geqq |(|x^+y| - |x^-y|)| = |(x^+ - x^-)y'| = |xy'| = x'y'$, et $|xy| \leqq |x^+y| + |x^-y| = (x^+ + x^-)y' = x'y'$, donc $|xy| = x'y' = |x| \cdot |y|$. De plus $xx' = x'^2$; or $xx' = (x^+)^2 - (x^-)^2$, donc $xx' \leqq (x^+)^2 + (x^-)^2$; mais $x'^2 = (x^+)^2 + (x^-)^2 + 2x^+x^-$, ce qui montre que $x^+x^- = 0$; on en tire que, pour x quelconque, $x^2 = (x^+)^2 + (x^-)^2 \geq 0$.

Soit alors $X \in F_+$; quels que soient λ et μ réels, et x et y quelconques dans E, on a $X((\lambda x + \mu y)^2) \geq 0$, d'où aussitôt l'inégalité de Schwartz

$$|X(xy)| \leq (X(x^2))^{\frac{1}{2}}(X(y^2))^{\frac{1}{2}} \tag{3}$$

et, comme conséquence immédiate, l'inégalité de Minkowski

$$(X((x + y)^2))^{\frac{1}{2}} \leq (X(x^2))^{\frac{1}{2}} + (X(y^2))^{\frac{1}{2}} \tag{4}$$

3. Considérons à présent une fonction linéaire positive U sur E, sur laquelle nous ferons d'abord les hypothèses suivantes:

(V) La relation $U(|x|) = 0$ entraîne $x = 0$.

(VI) Quel que soit $y \in E$, il existe un nombre fini $\|y\| \geq 0$ tel que, pour tout $x \in E$,

$$|U(yx)| \leq \|y\| \cdot U(|x|) \tag{5}$$

La première condition exprime que $U(|x|)$ est une *norme* sur E; la seconde, que, si on désigne par U_y la fonction linéaire $x \to U(yx)$, U_y est *continue* lorsqu'on topologise E par la norme $U(|x|)$.

Pour énoncer notre dernière hypothèse, désignons par S l'ensemble des $y \in E_+$ tels que, pour *tout* $x \in E_+$, on ait $yx \leq x$. Notre hypothèse est la suivante:

(VII) Quels que soient $X \in F$ tel que $0 \leq X \leq U$, et $x \in E$, on a

$$X(|x|) = \sup_{|y| \in S} |X(yx)| \tag{6}$$

On en déduit que, si $U(xy) = 0$ quel que soit $y \geq 0$, on a $x = 0$; en effet, on a alors, pour tout $y \in E$, $|U(xy)| \leq |U(xy^+)| + |U(xy^-)| = 0$, d'où $U(xy) = 0$; d'après (6) on en tire $U(|x|) = 0$, et cela entraîne $x = 0$ d'après (V). Il en résulte en particulier que, si $x^2 = 0$, on a $x = 0$, en vertu de l'inégalité de Schwartz.

4. Soit F_U l'ensemble des U_y, lorsque y parcourt E; d'après ce qui précède, l'application $y \to U_y$ est une application *biunivoque* de E sur F_U. Elle est évidemment *linéaire*; en outre, c'est un *isomorphisme* de la structure d'*ordre* de E sur celle de F_U. En effet, il est clair que $y' \leq y''$ entraîne $U_{y'} \leq U_{y''}$; mais réciproquement, $U_{y'} \leq U_{y''}$ entraîne $y' \leq y''$: il suffit de montrer que $U(yx) \geq 0$ quel que soit $x \geq 0$ entraîne $y \geq 0$. Or cette hypothèse entraîne en particulier $U(yy^-) \geq 0$, c'est-à-dire $U(-(y^-)^2) \geq 0$; comme $(y^-)^2 \geq 0$, ce n'est possible que si $U((y^-)^2) = 0$, d'où $(y^-)^2 = 0$, et par suite $y^- = 0$, c'est-à-dire $y \geq 0$.

A la topologie de F_U correspond sur E celle définie par les pseudo-normes $N'_y(x) = U_y(|x|)$ pour $y \geq 0$; cela résulte de l'identité $|U_z| = U_{|z|}$, que nous allons démontrer en nous appuyant sur l'hypothèse (VII), la formule (2), et un raisonnement emprunté au mémoire de N. Bourbaki. Il suffit évidemment de voir que, pour tout $x \geq 0$, $U(|z|x) \leq \sup \sum_i |U(zx_i)|$ pour toutes les partitions (x_i) de x. Or, d'après (VII), on a $U(|z|x) = \sup_{|y| \in S} |U(zxy)|$; mais

$$|U(zxy)| \leq |U(zxy^+)| + |U(zxy^-)| \leq \sum_i |U(zx_i)|$$

où (x_i) est la partition de x formée des trois éléments $x_1 = xy^+$, $x_2 = xy^-$, $x_3 = x - x|y|$; d'où la proposition.

L'application $y \to U_y$ se prolonge donc par continuité, en une application biunivoque du *complété* E_U de E, muni de la structure uniforme définie par les pseudo-normes N'_y, sur l'*adhérence* $\bar{F}_U$ de F_U dans F. Pour tout $y \in E_U$, on désignera encore par U_y l'élément de $\bar{F}_U$ qui lui correspond; on *ordonnera* E_U en y transportant la structure d'ordre de $\bar{F}_U$; cette structure transportée prolonge celle de E, d'après ce qui précède; l'ensemble E_U^+ des $y \geq 0$ est l'*adhérence* de E_+ dans E_U. En effet, il suffit de voir que, si G_U désigne l'ensemble des U_y tels que $y \in E_+$, $\bar{G}_U = \bar{F}_U \cap F_+$; or, l'ensemble $\bar{F}_U \cap F_+$, étant fermé, contient $\bar{G}_U$; et d'autre part, si $X \geq 0$ appartient à $\bar{F}_U$, à tout $x \in E_+$ et tout $\epsilon > 0$ correspond un $y \in E$ tel que $|X - U_y|(x) \leq \epsilon$, d'où, puisque $|X| = X$, et $|U_y| = U_{|y|}$, $|X - U_{|y|}|(x) \leq \epsilon$, ce qui montre que $X \in \bar{G}_U$. On en conclut aisément que, si $X \in \bar{F}_U$, $|X| \in \bar{G}_U$, d'après la continuité de la fonction $|X|$ dans F.

Il est clair par ailleurs que la structure d'espace vectoriel de E se prolonge à E_U, en une structure isomorphe à celle de $\bar{F}_U$; cette structure et la structure d'ordre de E_U satisfont aux axiomes (I), (II) et (III). En outre, *tout ensemble majoré dans E_U admet une borne supérieure*, car la borne supérieure, dans F, d'une partie majorée de $\bar{F}_U$ appartient à $\bar{F}_U$: c'est immédiat pour une partie *finie* $\left(\text{d'après la formule } \sup(X, Y) = \dfrac{X + Y + |X - Y|}{2}\right)$, et il en résulte que la borne supérieure d'une partie majorée quelconque de $\bar{F}_U$ est la limite d'un ensemble filtrant à droite sur $\bar{F}_U$, donc appartient à $\bar{F}_U$ (qui est fermé).

5. Pour une valeur *fixe* de $x \in E$, l'application linéaire $y \to yx$ de E dans E est *continue*; en effet, d'après la condition (VI), on a, pour tout $z \in E_+$, $N'_z(yx) = U(|yx|z) \leq \|x\| \cdot U(|y|z) = \|x\| \cdot N'_z(y)$.

On peut donc *prolonger* par continuité cette fonction à E_U ; on a ainsi défini le produit yx pour tout $y \in E_U$ et tout $x \in E$ (le produit yy', pour deux éléments de E_U, n'est pas défini en général!) Si x et x' sont deux éléments quelconques de E, on a $(yx)x' = y(xx')$. En effet, les deux membres sont des fonctions continues de y, égales dans E, qui est partout dense dans E_U, donc elles sont encore égales dans E_U. On montre de même les propriétés de distributivité $(y + y')x = yx + y'x$, $y(x + x') = yx + yx'$, et que, pour λ réel, $(\lambda y)x = y(\lambda x) = \lambda(yx)$. Si on remarque que la fonction sup (y, z) est uniformément continue dans $E_U \times E_U$, on démontre de la même manière que sup $(y, y')\cdot x =$ sup $(yx, y'x)$, et que $y\cdot$sup $(x, x') =$ sup (yx, yx') en prolongeant l'identité (IV).

Enfin, les fonctions $U(yx)$ et $U_y(x)$, pour x fixe dans E, sont linéaires en y, continues dans E et identiques; elles sont donc encore identiques pour tout $y \in E_U$.

6. Nous sommes maintenant en mesure d'énoncer le résultat final auquel nous voulons parvenir. Désignons par K ce que F. Riesz appelle "la plus petite famille complète contenant U"; cette famille est une partie de F_+ qui s'obtient de la manière suivante (F. Riesz, loc. cit., §V): on considère l'ensemble K' des fonctions linéaires positives X telles que $X \leq \lambda U$ pour une valeur convenable de $\lambda > 0$; K est l'ensemble des bornes supérieures de toutes les parties majorées de K'. Nous allons montrer que *l'ensemble K est identique à l'ensemble $\bar{G}_U$*. Toute fonction $X \in F_+$ se mettant, d'après le théorème fondamental de F. Riesz (loc. cit., th. 14) sous la forme $Y + Z$, où $Y \in K$ et où Z est *disjointe* de U, on aura $Y = U_y$ avec $y \in E_U$, d'après ce qui précède, et on obtiendra ainsi l'énoncé généralisant entièrement le théorème de Lebesgue-Nikodym aux fonctions linéaires abstraites.

Nous allons commencer par montrer que $K \subset \bar{G}_U$; comme toute partie majorée de $\bar{G}_U$ admet une borne supérieure appartenant à $\bar{G}_U$, il nous suffira de voir que $K' \subset \bar{G}_U$. Autrement dit, nous allons démontrer que, *si* $0 \leq X \leq \lambda U$ $(\lambda > 0)$, *il existe un $y \in E_U$ tel que $X = U_y$*.

Remarquons pour cela que, d'après (V) et les inégalités (3) et (4), $(U(x^2))^{\frac{1}{2}}$ est une *norme* dans E, et que, d'après (3), la fonction $(y, x) = U(yx)$ possède les propriétés d'un *produit scalaire*. Il en résulte aussitôt que, si on *complète* E, topologisé par la norme $(U(x^2))^{\frac{1}{2}}$, on obtient un *espace de Hilbert*[9] H. Comme, d'après (3), on a, pour tout $y \geq 0$,

$$N'_y(x) \leq (U(y^2))^{\frac{1}{2}}(U(x^2))^{\frac{1}{2}}$$

la topologie définie sur E par la norme $(U(x^2))^{\frac{1}{2}}$ est *plus fine* que celle définie par les pseudo-normes $N'_y(x)$; il s'ensuit que l'application identique φ de E, considéré comme sous-espace de H, sur E, considéré comme sous-espace de E_U,

[9] Par "espace de Hilbert," nous entendons un espace satisfaisant aux axiomes A, B, E énoncés au chap. I du mémoire de J. von Neumann, "*Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*," Math. Ann., *102* (1930), p. 49. Aucune hypothèse n'est faite sur le nombre de dimensions de l'espace, qui peut être, éventuellement, fini ou supérieur au dénombrable.

se prolonge en une application linéaire *continue* de H dans E_U, application que nous désignerons encore par φ. Remarquons maintenant que *toute* forme linéaire continue sur H est un produit scalaire (y, x), avec y fixe dans H. Or, pour tout x fixe dans E, (y, x) est une fonction continue de y dans H, qui coincide avec $U(yx) = U(\varphi(y)x)$ lorsque $y \in E$; mais $U(yx)$, considérée comme fonction de y, est continue dans E_U, donc $U(\varphi(y)x)$ est continue dans H. Par suite, (y, x) et $U(\varphi(y)x)$, qui coïncident dans la partie partout dense E de H, coïncident dans H.

Ceci posé, l'hypothèse $X \leq \lambda U$ entraîne, pour tout $y \in E$, $| X(yx) | \leq \lambda U(| y | \cdot | x |) = \lambda N'_{|y|}(x)$, autrement dit, l'application $x \to X(yx)$, que nous désignerons par X_y, est *continue* dans E, muni de la structure définie par les pseudo-normes N'_y ; à fortiori, elle est continue dans E, considéré comme sous-espace de l'espace de Hilbert H. Elle se prolonge donc dans H, et par suite, il existe $z_y \in E_U$ tel que $X_y = U_{z_y}$.

Il est immédiat que l'application $y \to z_y$ de E dans E_U est linéaire. Si $y \geq 0$, on a $z_y \geq 0$, car alors $U_{z_y}(x) = X_y(x) \geq 0$ pour tout $x \geq 0$. Enfin, z_y est fonction *continue* de y; en effet, pour tout $x \in E_+$ on a $N'_x(z_y) = U(| z_y | x) = U_{|z_y|}(x) = | U_{z_y} | (x) = | X_y | (x) = X_{|y|}(x) \leq \lambda N'_x(y)$ d'après l'identité démontrée au §4.

Faisons alors décrire à y l'ensemble S (§3); d'après la condition (IV), cet ensemble ordonné est filtrant à droite; il est *majoré* dans E_U, car $U_y \leq U$ quel que soit $y \in S$; son filtre des sections $\mathfrak{F}$ a donc une limite e dans E_U. En vertu de la continuité de z_y, cette fonction tend vers une limite z_e suivant le filtre $\mathfrak{F}$; pour tout x fixe dans E_+, d'après la continuité de $U(yx)$ dans E_U, on a donc $U(z_e x) = \lim_{\mathfrak{F}} U(z_y x) = \lim_{\mathfrak{F}} X(yx)$; mais, d'après (VII), $\lim_{\mathfrak{F}} X(yx) = X(x)$, ce qui démontre la première partie du théorème.

7. Il nous reste à voir que $\bar{G}_U \subset K$; d'après le théorème de décomposition de F. Riesz, tout $X \in \bar{G}_U$ peut se mettre d'une manière et d'une seule sous la forme $Y + Z$, où $Y \in K$ et où Z est disjointe de U; comme $K \subset \bar{G}_U$, $Z = X - Y$ appartient à $\bar{F}_U$, et comme $Z \geq 0$, $Z \in \bar{G}_U$; la proposition sera démontrée si on fait voir que toute fonction $Z \in \bar{G}_U$, disjointe de U, est nécessairement 0.

Nous utiliserons pour cela le critère suivant: si une fonction X appartient à $\bar{G}_U$, à tout $y \in E_+$ et tout $\epsilon > 0$ correspond un nombre $\eta > 0$ tel que les conditions $x \in E_+$, $x \leq y$, $U(x) \leq \eta$ entraînent $| X(x) | \leq \epsilon$ (critère d'"absolue continuité"; nous en empruntons l'énoncé et la démonstration au travail de N. Bourbaki). En effet, il existe par hypothèse $z \in E$ tel que $| X - U_z | (y) \leq \delta$, d'où, pour la partition $(x, y - x)$ de y, $| X(x) - U(zx) | + | X(y - x) - U(z(y - x)) | \leq \delta$ et a fortiori $| X(x) - U(zx) | \leq \delta$; par suite $| X(x) | \leq \delta + \| z \| \cdot U(x)$, ce qui est aussi petit qu'on veut si δ, puis $U(x)$, sont pris assez petits.

Si Z appartient à $\bar{G}_U$ et est disjointe de U, on a par définition inf $(Z, U) = 0$; et, pour tout $x \in E_+$ et tout $\eta \geq 0$, il existe une partition (y, z) de x telle que $U(y) + Z(z) \leq \eta$,[10] d'où en particulier $U(y) \leq \eta$; si $\eta \leq \epsilon$ est choisi de sorte

[10] *F. Riesz*, loc. cit., p. 183.

que les conditions $y \leq x$, $U(y) \leq \eta$ entraînent $Z(y) \leq \epsilon$, on aura donc $Z(x) = Z(y) + Z(z) \leq 2\epsilon$, et comme ϵ est arbitraire, $Z(x) = 0$, d'où la proposition, qui achève de démontrer le théorème.

Le même raisonnement montre aussi que la condition donnée ci-dessus pour que $X \in \bar{G}_U$, est non seulement *nécessaire*, mais aussi *suffisante*.

8. Si les éléments de E sont des *fonctions réelles finies*, définies sur un ensemble A, les axiomes (II) et (III) sont vérifiés d'eux-mêmes, et (I) entraîne (IV) (on suppose bien entendu que l'addition et la multiplication sont l'addition et la multiplication ordinaires des fonctions réelles). L'axiome (VI) est vérifié pour *toute* fonction linéaire positive U si les éléments de E sont des fonctions *bornées* sur A.

Quant à la condition (VII), N. Bourbaki a montré, dans le travail auquel nous nous sommes déjà plusieurs fois référé, qu'elle est vérifiée dans chacun des deux cas suivants: 1° E contient la fonction égale à la constante $+1$; 2° si $x \in E_+$, $\sqrt{x} \in E$. Il est vraisemblable que ces deux cas ne sont pas les seuls où la condition (VII) est réalisée; mais nous allons voir par contre, sur un exemple, que cette condition n'est pas une conséquence des autres, et que, lorsqu'elle n'est pas vérifiée, le théorème de Lebesgue-Nikodym peut se trouver en défaut.

Prenons pour A l'intervalle $0 \leq t \leq 1$, et pour E l'ensemble des fonctions réelles *continues* dans A, *nulles* au point $t = 0$, et *dérivables* en ce point. Il est immédiat que E est un espace vectoriel, et qu'il satisfait aux conditions (I), (II) et (III); en outre le produit de deux fonctions de E appartient à E, et la condition (IV) est évidemment vérifiée. Les fonctions de E sont bornées, donc (VI) est satisfaite pour toute fonction linéaire positive.

Prenons alors $U(x) = x'(0) + \int_0^1 x(t)\, dt$; il est clair que (V) est vérifiée. Par contre, nous allons voir que (VII) ne l'est pas. En effet, pour tout $y \in E$, on a $U(yx) = \int_0^1 y(t)x(t)\, dt$; si $|y(t)| \leq 1$ quel que soit $t \in A$, on a $|U(yx)| \leq \int_0^1 |x(t)|\, dt$, donc, pour tout $x \geq 0$, $|U(x) - U(yx)| \geq x'(0)$, et si $x'(0) \neq 0$, (VII) ne peut être satisfaite.

Pour montrer que, dans ce cas, le théorème de Lebesgue-Nikodym, tel que nous l'avons énoncé, n'est pas vrai, nous établirons que *U n'appartient pas à $\bar{G}_U$*. Sinon, en effet, pour tout $x \in E_+$ et tout $\epsilon > 0$, il existerait $y \in E_+$ tel que $|U - U_y|(x) \leq \epsilon$, c'est-à-dire, pour toute partition (x_i) de x, $\sum_i |U(x_i) - U(yx_i)| \leq \epsilon$, et en particulier, $|x_i'(0) + \int_0^1 (1 - y(t))x_i(t)\, dt| \leq \epsilon$ pour tout indice i. Or, on peut choisir la partition (x_i) de sorte que, pour un indice k, $x_k'(0) = x'(0)$, et $|\int_0^1 (1 - y(t))x_k(t)\, dt| \leq \epsilon$. On aurait donc $|x'(0)| \leq 2\epsilon$, quel que soit $\epsilon > 0$, ce qui est absurde, puisque x a été prise arbitraire dans E.

NANCY, FRANCE.

ANNALS OF MATHEMATICS
Vol. 42, No. 2, April, 1941

ON THE MODULAR CHARACTERS OF GROUPS

BY R. BRAUER AND C. NESBITT*

(Received, December 15, 1939)

PART I. Introduction

§1. Ordinary representations. Group ring. §2. Arithmetical questions. §3. Modular representations. §4. Decomposition numbers. §5. Cartan invariants. §6. Characters. §7. The character relations. §8. Corollaries. §9. Blocks. §10. Decomposition of $\bar{\Gamma}$. §11. Summary of the results.

PART II. Blocks of highest kind

§12. Condition for the reducibility of $\bar{Z}_i$. §13. Blocks of highest kind. §14. Vanishing of the character for p-singular elements of $\mathfrak{G}$. §15. Example.

PART III. The elementary divisors of the Cartan matrix

§16. Computation of the elementary divisors of C. §17. Blocks of type α. §18. Ordinary characters which are linearly independent (mod $\mathfrak{p}$). §19. Application of the lemma of §18. §20. Blocks of the lowest kind. §21. Alternative proof of theorem 2.

PART IV. On the multiplication of characters

§22. Relations between the problems of determining the ordinary and the modular characters of $\mathfrak{G}$. §23. The multiplication of characters. §24. Upper and lower bounds for the degrees of the indecomposable constituents U_κ.

PART V. Relations between the characters of a group $\mathfrak{G}$ and those of a subgroup $\mathfrak{H}$ of $\mathfrak{G}$

§25. The induced character. §26. The formulas of Nakayama. §27. On the converse of theorem 1. §28. An upper and a lower bound for c_{11}.

PART VI. Special cases and examples

§29. Special cases. §30. The groups $GLH(2, p^a)$, $SLH(2, p^a)$, and $LF(2, p^a)$. §31. The Cartan invariants and decomposition numbers (for p) of $LF(2, p)$.

---

* Presented to the American Mathematical Society, October 30, 1937.

I. Introduction[1]

1. Ordinary Representations. Group ring. The representations of a group $\mathfrak{G}$ of finite order g were first treated by *Frobenius*[2] in his theory of group characters. The coefficients of the linear transformations are taken as complex numbers, but it does not make any difference if we take them as the elements of an algebraically closed field K of characteristic 0. The theory has been extended by *I. Schur*[3] to the case where K is any field of characteristic 0. It does not mean an essential restriction, if we take K as an algebraic number field.

Instead of considering representations of $\mathfrak{G}$, we may consider representations of the group ring[4] Γ of $\mathfrak{G}$ with regard to K. This Γ is an associative algebra consisting of all symbols

$$\alpha = a_1G_1 + a_2G_2 + \cdots + a_gG_g \tag{1}$$

where $G_1, G_2, \cdots, G_g$ are the elements of $\mathfrak{G}$, and $a_1, a_2, \cdots, a_g$ are arbitrary elements of K. The equality of two such elements, their addition, and their multiplication are defined in a natural manner. The study of the representations of Γ is closely tied up with the investigation of the algebra Γ.

2. Arithmetical questions. We may also study Γ from an arithmetical point of view. Taking K as an algebraic number field, we obtain a domain of integrity $\mathfrak{J}$ if we take the a_i in (1) from the domain $\mathfrak{o}$ of the integers of K. The question arises in what manner does a prime ideal $\mathfrak{p}$ behave when considered as an ideal of $\mathfrak{J}$. The behavior of $\mathfrak{p}$ in $\mathfrak{J}$ is characterized by the structure of the residue class ring $\mathfrak{J}/\mathfrak{p}$. This ring can be considered as an algebra $\bar{\Gamma}$ over the residue class field $\bar{K} = \mathfrak{o}/\mathfrak{p}$ of the integers of K taken (mod $\mathfrak{p}$). Obviously, $\bar{\Gamma}$ is the group ring of $\mathfrak{G}$ with regard to the finite ground field $\bar{K}$. The study of the structure of $\bar{\Gamma}$ then amounts essentially to the same thing as the study of the representations of $\bar{\Gamma}$ or of $\mathfrak{G}$ by matrices with coefficients in the finite field $\bar{K}$. We thus are led to the problem of extending Frobenius' theory to the case of a modular field of reference (i.e. a field of a characteristic $p \neq 0$).

3. Modular representations. Modular representations of a group $\mathfrak{G}$ (i.e. representations of $\mathfrak{G}$ by matrices with coefficients in a modular field) were first

[1] In §§4–10 of the introduction, we give a short account of the theory of modular representations of a group as developed in our paper: On the modular representations of groups of finite order, University of Toronto Studies, Math. Series No. 4, 1937 (we refer to this paper as M.R.). We tried to make it unnecessary for a reader, who is familiar with the theory of representations in general, to read our former paper. An exception is formed perhaps by the proof of formula (5) below, but literature for other proofs of this formula are mentioned in footnote 10.

[2] For Frobenius' theory, see the accounts in L. E. Dickson, *Modern Algebraic Theories*, Chicago, 1926, chapter XIV; G. A. Miller, H. F. Blichfeldt, L. E. Dickson, *Theory and Application of Finite Groups*, New York 1916, chapter XIII, H. F. Blichfeldt, *Finite Collineation Groups*, Chicago 1917, chapter VI.

[3] I. Schur, Sitzungsber. Preuss. Akad., 1906, p. 164.

[4] Cf., for instance, H. Weyl, *The Classical Groups*, Princeton 1939, Chapter III.

studied by Dickson.[5] He proved that Frobenius' theory remains valid, if the characteristic p of the field is prime to the order g of $\mathfrak{G}$. Since the discriminant of Γ is a power of g, this corresponds to the case that the prime ideal $\mathfrak{p}$ in §2 is not a discriminant divisor. If, however, p divides g, then we must expect results which differ from those of Frobenius. This was shown first by a theorem of Dickson[6] concerning the splitting of the regular representation (cf. §8 below). A coherent theory of the modular representations was given by the authors in a previous paper.[7] In the following §§4–9, we shall discuss briefly our former results. We prefer, in most of what follows, to use the language of the theory of representations (instead of that of the theory of algebras or of the theory of ideals).

4. Decomposition numbers. We choose the algebraic number field K such that the absolutely irreducible representations of $\mathfrak{G}$ in the sense of Frobenius can be written with coefficients in K. Let $Z_1, Z_2, \cdots, Z_n$ be the essentially different ones among these representations, and let z_i denote the degree of Z_i. Then n is the number of classes of conjugate elements $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_n$ in $\mathfrak{G}$.

Let p be a fixed rational prime number, and $\mathfrak{p}$ be a fixed prime ideal divisor of p in K. We may assume that the coefficients of all the Z_i are $\mathfrak{p}$-integers (i.e. numbers of the form α/β where α and β are integers of K, and β is prime to $\mathfrak{p}$). Let $\mathfrak{o}_\mathfrak{p}$ be the ring of the $\mathfrak{p}$-integers of K, and $\bar{K}$ the residue class field of $\mathfrak{o}_\mathfrak{p}$ (mod $\mathfrak{p}$) which is identical with the field $\mathfrak{o}/\mathfrak{p}$ in §2. We denote generally the residue class of an element z of K (mod $\mathfrak{p}$) by $\bar{z}$. Similarly, replacing every coefficient z in a representation Z of $\mathfrak{G}$ with coefficients in $\mathfrak{o}_\mathfrak{p}$ by its residue class $\bar{z}$, we obtain a modular representation $\bar{Z}$ with coefficients in $\bar{K}$. In this manner we may form $\bar{Z}_1, \bar{Z}_2, \cdots, \bar{Z}_n$. These modular representations will, in general, be reducible and will then split into irreducible modular representations F_κ with coefficients in $\bar{K}$. We indicate by

$$\bar{Z}_i \leftrightarrow \sum_\kappa d_{i\kappa} F_\kappa \tag{2}$$

that F_κ will appear in $\bar{Z}_i$ with some multiplicity $d_{i\kappa}$. These rational integers $d_{i\kappa} \geqq 0$, and are called the "decomposition numbers" of $\mathfrak{G}$. In the sense of §2, they describe a connection between the simple invariant subalgebras of Γ, and the prime ideal divisors of $\mathfrak{p}$ in $\mathfrak{J}$.

5. Cartan invariants. Of special importance is the regular representation $\bar{R}$ of $\mathfrak{G}$ (or $\bar{\Gamma}$) formed with regard to $\bar{K}$ as ground field. Since the group ring is no longer semisimple in the modular case, the theorem of the full reducibility of $\bar{R}$ does not hold any more. Let $U_1, U_2, \cdots, U_k$ be the distinct indecom-

[5] L. E. Dickson, Transact. Am. Math. Soc. 8, 1907, p. 389.

[6] L. E. Dickson, Bull. Amer. Math. Soc. 13, 1907, p. 477.

[7] For the following, cf. M.R., and also R. Brauer, Nat. Ac. of Sciences 25, 1939, p. 252. We refer to this last paper as R.A.

posable constituents of $\bar{R}$. Each U_κ can still be broken up into its irreducible constituents in $\bar{K}$. This splitting is of the form

$$U_\kappa = \begin{pmatrix} F_\kappa & & & \\ & F_* & & \\ & & \ddots & \\ & * & & F_\kappa \end{pmatrix} \tag{3}$$

if the notation is chosen suitably.[8] The representations $F_1, F_2, \cdots, F_k$ are all distinct, and there are no other irreducible representations of $\mathfrak{G}$ in $\bar{K}$. Further, the F_κ are absolutely irreducible.

We denote the degree of F_κ by f_κ, that of U_κ by u_κ. Then U_κ appears f_κ times as indecomposable constituent of $\bar{R}$ and F_κ appears u_κ times as irreducible constituent of $\bar{R}$.

Let $c_{\kappa\lambda}$ be the multiplicity of F_λ as irreducible constituent of U_κ

$$U_\kappa \leftrightarrow \sum_\lambda c_{\kappa\lambda} F_\lambda. \tag{4}$$

Here, the $c_{\kappa\lambda}$ are rational integers $\geqq 0$, the Cartan invariants[9] of $\mathfrak{G}$ (for p). They also can be characterized by means of structural properties of Γ; they express mutual relations between the different prime ideal divisors of $\mathfrak{p}$ in $\mathfrak{J}$. Between the decomposition numbers and the Cartan invariants, we have the following equations[10]

$$c_{\kappa\lambda} = \sum_{i=1}^{n} d_{i\kappa} d_{i\lambda} \qquad (\kappa, \lambda = 1, 2, \cdots, k) \tag{5}$$

or in matrix form

$$C = D'D \tag{6}$$

where $C = (c_{\kappa\lambda})$, $D = (d_{i\kappa})$ and D' is the transpose of D.

There exists a representation (U_κ) of $\mathfrak{G}$ in K which if taken (mod $\mathfrak{p}$) becomes similar to U_κ, $\overline{(U_\kappa)} = U_\kappa$. We then have[11]

$$(U_\kappa) \leftrightarrow \sum_i d_{i\kappa} Z_i. \tag{7}$$

6. Characters. Let M be a representation of $\mathfrak{G}$ which represents the group element G by $M(G)$. We denote the trace of the matrix $M(G)$ by $\chi(G)$. Then $\chi(G)$ is a function of the arbitrary group element G, the *character* of M. The

---

[8] See R. Brauer and C. Nesbitt, Nat. Ac. of Sciences 23, 1937, p. 236; C. Nesbitt, Ann. of Math. 39, 1938, p. 634. Free places in matrices are to be replaced by 0, the * stand for quantities in which we are not interested.

[9] E. Cartan, Annales de Toulouse 12B, 1898, p. 1.

[10] Three different proofs are given in M.R. pp. 9–11; T. Nakayama, Ann. of Math. 39, 1938, p. 361; R.A. pp. 257–258.

[11] Cf. R. A. The use of this fact which has not been mentioned in M.R. can be avoided, see footnote 13.

value of χ is the same for conjugate elements of $\mathfrak{G}$. We may, therefore, consider χ as a function of the classes of conjugate elements $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_n$; we set $\chi_\nu = \chi(\mathfrak{C}_\nu) = \chi(G)$ if G belongs to $\mathfrak{C}_\nu$.

Let N be a second representation of $\mathfrak{G}$ with the character φ. We consider first the case that we have a ground field K of characteristic 0. If the determinants of $M(G)$ and $N(G)$ do not vanish, then the characters are equal, $\chi = \varphi$, if and only if M and N have the same irreducible constituents. Because of the full reducibility, this is the same as similarity of M and N. If we admit matrices of determinant 0 in M and N, we must add the assumption that M and N have the same degree. Otherwise, the (0)—representation may appear with different multiplicities in M and N.[12]

In the case of a ground field K of characteristic p, these theorems are not true. However, the method by which they are proved allows one to show that M and N have the same irreducible constituents, if and only if $M(G)$ and $N(G)$ have the same characteristic roots for every G in $\mathfrak{G}$.

We may write G as a product AB of two commutative elements where A has an order prime to p, whereas B has an order p^β, $\beta \geqq 0$. The characteristic roots of $M(B)$ are all 1, being p^β-th roots of unity in a field of characteristic p. It follows that $M(G)$ and $M(A)$ have the same characteristic roots. It is, therefore, sufficient to require above that $M(G)$ and $N(G)$ have the same characteristic roots for every G of an order prime to p. Then the same will be automatically true for all G in $\mathfrak{G}$. We call an element G of $\mathfrak{G}$ *p-regular* if its order is prime to p.

We use the same notations as in §2. We set

$$g = p^\alpha \cdot g' \qquad (p, g') = 1. \tag{8}$$

Let K_1 be the field obtained from K by the adjunction of the g'-th roots of unity $1, \delta, \delta^2, \cdots, \delta^{g'-1}$, let $\mathfrak{p}_1$ be a prime ideal divisor of $\mathfrak{p}$ in K_1, and let $\bar{K}_1$ be the field of integers of K_1 taken mod $\mathfrak{p}_1$. Then $\bar{K}_1$ is an extension field of $\bar{K}$, which contains the modular g'-th roots of unity $1, \bar{\delta}, \bar{\delta}^2, \cdots, \bar{\delta}^{g'-1}$, the residue classes of $1, \delta, \cdots \delta, \delta^{g'-1}$ (mod $\mathfrak{p}_1$). We have a $(1 - 1)$ relation between the ordinary and the modular g'-th roots of unity since $\delta^\alpha \not\equiv \delta^\beta$ (mod $\mathfrak{p}_1$) if $\delta^\alpha \neq \delta^\beta$.

If F now is a modular representation of $\mathfrak{G}$ with coefficients in $\bar{K}$ or in an extension field of $\bar{K}$, the characteristic roots of $F(G)$ will lie in $\bar{K}_1$. Let G be a p-regular element of $\mathfrak{G}$. We replace each such root $\bar{\delta}^\nu$ by δ^ν, and define now $\chi(G)$ as the sum of these δ^ν. In this manner, the character $\chi(G)$ is defined as a complex number for the p-regular elements G; the original value was the residue class $\bar{\chi}(G)$ of $\chi(G)$ (mod $\mathfrak{p}$). It now follows easily that two modular representations (with coefficients in $\bar{K}$ or in an extension field of $\bar{K}$) have the same irreducible constituents if and only if the two characters in the new sense coincide for p-regular elements.

---

[12] G. Frobenius and I. Schur, Sitzungsber. Preuss. Akad. 1906, p. 1906, p. 209.

7. The character relations. Let $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_{k'}$ be the classes of conjugate elements which contain the p-regular elements. We denote by $\eta^{(\kappa)}$ the character of U_κ, by $\varphi^{(\kappa)}$ that of F_κ (cf. §5), by $\zeta^{(i)}$ that of Z_i. The value of a character for the class $\mathfrak{C}_\nu$ will be indicated by a suffix ν, e.g. $\eta_\nu^{(\kappa)}$, $(\nu = 1, 2, \cdots, k')$. The relations (7) and (2) now give[13]

$$\eta^{(\kappa)} = \sum_i d_{i\kappa}\zeta^{(i)} \tag{9}$$

$$\zeta^{(i)} = \sum_\lambda d_{i\lambda}\varphi^{(\lambda)} \tag{10}$$

$(i = 1, 2, \cdots, n; \kappa = 1, 2, \cdots, k)$. From these and (5), or directly from (4) we have

$$\eta^{(\kappa)} = \sum_\lambda c_{\kappa\lambda}\varphi^{(\lambda)}. \tag{11}$$

In particular, for the degrees u_κ, z_i, f_λ of U_κ, Z_i, F_λ, respectively, (9), (10), (11) for the unit element give

$$u_\kappa = \sum_i d_{i\kappa}z_i, \qquad z_i = \sum_\lambda d_{i\lambda}f_\lambda, \qquad u_\kappa = \sum_\lambda c_{\kappa\lambda}f_\lambda \tag{12}$$

since $u_\kappa = \eta^{(\kappa)}(1)$, $z_i = \zeta^{(i)}(1)$, $f_\lambda = \varphi^{(\lambda)}(1)$. We arrange $\varphi_\lambda^{(\kappa)}$, $\eta_\lambda^{(\kappa)}$ $\zeta_\lambda^{(i)}$ in matrix form

$$\Phi = (\varphi_\lambda^{(\kappa)}), \qquad \mathrm{H} = (\eta_\lambda^{(\kappa)}), \qquad \mathrm{Z} = (\zeta_\lambda^{(i)})$$

(κ row index, λ column index in Φ, H; i row index, λ column index in Z; $\kappa = 1, 2, \cdots, k$; $\lambda = 1, 2, \cdots, k'$; $i = 1, 2, \cdots, n$). Then relations (9), (10) and (11) become

$$\mathrm{H} = D'\mathrm{Z}, \qquad \mathrm{Z} = D\Phi, \qquad \mathrm{H} = C\Phi. \tag{13}$$

From the orthogonality relations for the ordinary group characters, we obtain

$$\mathrm{Z}'\mathrm{Z} = (g/g_\kappa\delta_{\kappa\lambda^*}) = \mathrm{T} \tag{14}$$

where g_κ denotes the number of elements in the class $\mathfrak{C}_\kappa$, and where the class $\mathfrak{C}_{\kappa^*}$ contains the elements reciprocal to those of $\mathfrak{C}_\kappa$ so that $1^*, 2^*, \cdots, k'^*$ is a permutation of $1, 2, \cdots, k'$. Then (14), (13) and (6) yield

$$\mathrm{H}'\Phi = \Phi'C\Phi = \mathrm{T}. \tag{15}$$

The equation (15) contains in matrix form orthogonality relations for the modular group characters, viz.

$$\sum_\rho \eta_\nu^{(\rho)}\varphi_\mu^{(\rho)} = \sum_{\rho,\sigma}\varphi_\nu^{(\rho)}c_{\rho\sigma}\varphi_\mu^{(\sigma)} = g\delta_{\nu\mu^*}/g_\nu \tag{16}$$

(see also relations (20), (21) and (22) below).

---

[13] We may avoid the use of (7) here by first deriving (10) from (2) and then (9) from (4), (5) and (10), see M.R.

8. Corollaries. Since T in (15) is non-singular, the columns of the matrix Φ which is of type (k, k') [14] are linearly independent, and hence $k \geqq k'$. On the other hand, the rows are linearly independent (mod $\mathfrak{p}$) because a linear relation would give a linear relation among the characters of $F_1, F_2, \cdots, F_k$, the values of these characters being understood as numbers of $\bar{K}$, as at the beginning of §6. Such a relation is impossible, hence $k = k'$. The number of distinct absolutely irreducible modular representations is equal to the number of classes of conjugate p-regular elements in $\mathfrak{G}$.[15] Further the determinant $| \Phi |$ of Φ is prime to $\mathfrak{p}$. Since $| \Phi |$ is integral, and, its square is rational according to (15) we see that $| \Phi |$ is prime to p.

$$(17) \qquad (| \Phi |, p) = 1.$$

The column of H corresponding to the unit element of $\mathfrak{G}$ consists of $u_1, u_2, \cdots, u_k$. Since here $g_\nu = 1$, we obtain from (16) and (17) Dickson's theorem:[16]

$$(18) \qquad u_\kappa \equiv 0 \quad (\text{mod } p^a), \qquad (\kappa = 1, 2, \cdots, k).$$

We have also now that all matrices which appear in (15) have inverses. Let us set, in particular, $C^{-1} = (\gamma_{\kappa\lambda})$. It follows from (15) that

$$(19) \qquad \Phi T^{-1} H' = (\delta_{\kappa\lambda})$$

which gives the character relations

$$(20) \qquad \sum_\nu g_\nu \varphi_\nu^{(\kappa)} \eta_{\nu^*}^{(\lambda)} = \delta_{\kappa\lambda} g \qquad (\kappa, \lambda = 1, 2, \cdots, k).$$

In addition, multiplying (19) through by $C^{-1} = (\gamma_{\kappa\lambda})$ and using (13), we have

$$(21) \qquad \Phi T^{-1} \Phi' = C^{-1}, \text{ that is, } \sum_\nu g_\nu \varphi_\nu^{(\kappa)} \varphi_{\nu^*}^{(\lambda)} = \gamma_{\kappa\lambda} g \qquad (\kappa, \lambda = 1, 2, \cdots, k)$$

and from (19) multiplied through by C

$$(22) \qquad HT^{-1}H' = C, \text{ that is, } \sum_\nu g_\nu \eta_\nu^{(\kappa)} \eta_{\nu^*}^{(\lambda)} = c_{\kappa\lambda} g \qquad (\kappa, \lambda = 1, 2, \cdots, k).$$

If $(p, g) = 1$, then we have full reducibility, $U_\kappa = F_\kappa$. The matrix C, and then also D, is equal to the unit matrix, and we have $\bar{Z}_i = F_i$ (Speiser[17]).

9. Blocks. It is well known and easy to prove that the n elements

$$(23) \qquad \Omega_\nu = \sum_{G \text{ in } \mathfrak{C}_\nu} G \qquad (\nu = 1, 2, \cdots, n)$$

form a basis of the centre of the group ring. Each irreducible representation of $\mathfrak{G}$ represents Ω_ν by a scalar multiple of the unit matrix I. We see

$$(24) \qquad Z_i(\Omega_\nu) = \omega_\nu^{(i)} I, \qquad F_\kappa(\Omega_\nu) = \psi_\nu^{(\kappa)} I$$

---

[14] By a matrix of type (a, b), we understand a matrix with a rows and b columns.

[15] See R. Brauer, Actual. Scient. 195, Paris, 1935; M.R.

[16] See footnote 6, also M.R.

[17] Cf. A. Speiser, Theorie der Gruppen von endlicher Ordnung, 3rd edition, Berlin 1937, p. 223.

where $\omega_\nu^{(i)}$ is an integer of K, and $\psi_\nu^{(\kappa)}$ lies in $\bar{K}$. We say that F_κ and F_λ belong to the same *block*, if $\psi_\nu^{(\kappa)} = \psi_\nu^{(\lambda)}$ for $\nu = 1, 2, \cdots, n$. Then F_κ and F_λ represent the centre of $\bar{\Gamma}$ essentially in the same manner. Thus $F_1, F_2, \cdots, F_k$ appear distributed into s "*blocks*" $B_1, B_2, \cdots, B_s$.

We also speak of the U_κ which belong to a block B_τ by counting U_κ in B_τ if F_κ belongs to B_τ. Each matrix $U_\kappa(\Omega_\nu)$ can have only one characteristic root[18] which necessarily is $\psi_\nu^{(\kappa)}$ since F_κ is a constituent of U_κ, cf. (3). It follows that all the irreducible constituents of U_κ belong to B_τ. More generally, if in the sequence

$$U_\kappa, U_\alpha, U_\beta, \cdots, U_\sigma, U_\lambda \tag{25}$$

any two consecutive U_ρ have an irreducible constituent in common, then all the U_ρ and their irreducible constituents belong to the same block B_τ. If however U_κ and U_λ cannot be joined by such a chain (25), then it is easy to construct a centre element of $\bar{\Gamma}$ which is represented by I in F_κ and by 0 in F_λ so that F_κ and F_λ do not belong to the same block. We have here a new characterization of the blocks.

Assume now that $\bar{Z}_i$ contains F_κ as a irreducible constituent. From (24) it follows that

$$\bar{\omega}_\nu^{(i)} = \psi_\nu^{(\kappa)}$$

where the bar again indicates the residue class (mod $\mathfrak{p}$). All the irreducible constituents of $\bar{Z}_i$ belong necessarily to the same block B_τ. We now say that Z_i also belongs to the block B_τ. Two ordinary representations Z_i and Z_j belong to the same block if and only if $\omega_\nu^{(i)} \equiv \omega_\nu^{(j)}$ (mod $\mathfrak{p}$) for $\nu = 1, 2, \cdots, n$. Comparing the trace in the first formula (24) in a well-known manner, we obtain

$$\omega_\nu^{(i)} = g_\nu \zeta_\nu^{(i)}/z_i \tag{26}$$

where z_i is the degree of Z_i. Hence, Z_i and Z_j belong to the same block if and only if

$$g_\nu \zeta_\nu^{(i)}/z_i \equiv g_\nu \zeta^{(j)}/z_j \pmod{\mathfrak{p}} \qquad (\nu = 1, 2, \cdots, n). \tag{27}$$

In what follows we shall always take $\varphi^{(1)}, \zeta^{(1)}$ to be the character of the unit representation considered as a modular and as an ordinary irreducible representation, respectively, of $\mathfrak{G}$, and B_1 to be the block which contains these characters.

We arrange the $F_1, F_2, \cdots, F_k$ and $Z_1, Z_2, \cdots, Z_n$ such that we first take the representations of B_1, then those of B_2, etc. Let x_τ be the number of Z_i belonging to B_τ and y_τ the number of F_i belonging to B_τ. It follows that C and D break up

$$C = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & C_s \end{pmatrix} \qquad D = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & D_s \end{pmatrix} \tag{28}$$

[18] Cf. R. Brauer and I. Schur, Sitzungsber. Preuss. Akad. 1930, p. 209, §2.

where C_τ is a square matrix of degree y_τ, and D_τ is of type (x_τ, y_τ). It is impossible to arrange the representations in such a manner that C_τ or D_τ break up further. For C_τ this follows directly from the properties of blocks given above. But because of (6), we have

$$C_\tau = D_\tau' D_\tau. \tag{29}$$

A breaking up of D_τ would imply one of C_τ. Since C_τ is non-singular (cf. (15) and (14)), we must have

$$x_\tau \geqq y_\tau. \tag{30}$$

We now form the trace of the element $F_\kappa(\Omega_\nu)$ in (24) and find $g_\nu \bar{\varphi}_\nu^{(\kappa)}$, since Ω_ν is the sum of g_ν elements all of which have the trace $\bar{\varphi}_\nu^{(\kappa)}$ in the representation F_κ. On the other hand, the trace of $F_\kappa(\Omega_\nu)$ is $f_\kappa \psi_\nu^{(\kappa)}$. If F_κ appears as modular constituent of the ordinary representation Z_i, then $\psi_\nu^{(\kappa)}$, as we have seen is the residue class of $\omega_\nu^{(i)}$ (mod $\mathfrak{p}$). Moreover, $\omega_\nu^{(i)}$ (mod $\mathfrak{p}$) depends only on the block B_τ to which F_κ and Z_i belong. We indicate that by setting $\omega_\nu^{(i)} \equiv \theta_\nu^{(\tau)}$ (mod $\mathfrak{p}$), where $\theta^{(\tau)}$ depends only on τ and not on i. Hence

$$g_\nu \varphi_\nu^{(\kappa)} \equiv f_\kappa \theta_\nu^{(\tau)} \pmod{\mathfrak{p}}. \tag{31}$$

Let us set $gC^{-1} = (\tilde{\gamma}_{\kappa\lambda})$, that is, $\tilde{\gamma}_{\kappa\lambda} = g\gamma_{\kappa\lambda}$. From (21) it follows that the $\tilde{\gamma}_{\kappa\lambda}$ are algebraic integers; that they are rational comes as a consequence of their definition. From (31) and (21) we have

$$\tilde{\gamma}_{\kappa\lambda} \equiv f_\kappa \sum_{\nu=1}^{k} \theta_\nu^{(\tau)} \varphi_{\nu^*}^{(\lambda)} \pmod{\mathfrak{p}} \tag{32}$$

The sum on the right depends on τ and λ; we denote it by $S(\tau, \lambda)$. If F_λ also belongs to the block B_τ, then by reasons of symmetry

$$\tilde{\gamma}_{\kappa\lambda} \equiv f_\kappa S(\tau, \lambda) \equiv f_\lambda S(\tau, \kappa) \pmod{\mathfrak{p}}. \tag{33}$$

In particular, if f_κ or f_λ are divisible by p, then $\tilde{\gamma}_{\kappa\lambda}$ is divisible by $\mathfrak{p}$ and hence $\tilde{\gamma}_{\kappa\lambda} \equiv 0 \pmod{p}$. If $f_\kappa \not\equiv 0 \pmod{p}$, then (33) shows that the value of $S(\tau, \kappa)/f_\kappa$ (mod $\mathfrak{p}$) depends only on τ. We may, therefore, write

$$\tilde{\gamma}_{\kappa\lambda} \equiv f_\kappa f_\lambda S_\tau \pmod{\mathfrak{p}}$$

where S_τ depends only on τ. We may here take S_τ as a rational integer and have

$$\tilde{\gamma}_{\kappa\lambda} \equiv f_\kappa f_\lambda S_\tau \pmod{p} \qquad \text{if } F_\kappa \text{ and } F_\lambda \text{ in } B_\tau. \tag{34}$$

If F_κ and F_λ belong to different blocks, then $\tilde{\gamma}_{\kappa\lambda} = 0$, because of the form (28) of C.

Since $\varphi^{(1)}$ is the character of the 1-representation and is contained in the block B_1, then (21) and (34) show

$$N = \tilde{\gamma}_{11} \equiv S_1 \pmod{p} \tag{35}$$

where N is the number of all p-regular elements of $\mathfrak{G}$.

10. Decomposition of Γ. The block properties derived in §9 are those which we are going to use in the later sections. But the importance of these blocks can better be recognized from other facts which we shall describe briefly.[19] Let B_τ be a fixed block, and consider the elements $\bar{\alpha}$ of $\bar{\Gamma}$, for which $U_\kappa(\bar{\alpha}) = 0$ for every U_κ except for those of B_τ. These $\bar{\alpha}$ form an invariant subalgebra $\sum_\tau$, and we have

$$\bar{\Gamma} = \sum\nolimits_1 \oplus \sum\nolimits_2 \oplus \cdots \oplus \sum\nolimits_s .$$

The $\sum_\tau$ cannot be represented as direct sums.

In close connection with this fact, we have the following ideal theoretical significance of the blocks. The ideal $\mathfrak{p}$ of $\mathfrak{J}$ (cf. §2) can be written uniquely as the intersection of s ideals $\mathfrak{M}_\tau$ ($\neq \mathfrak{J}$) any two of which are relatively prime

$$\mathfrak{p} = [\mathfrak{M}_1, \mathfrak{M}_2, \cdots, \mathfrak{M}_s]$$

and $\mathfrak{M}_\tau$ cannot be written as intersection of relatively prime ideals $\neq \mathfrak{M}_\tau$. There are exactly k prime ideal divisors $\mathfrak{P}_1, \mathfrak{P}_2, \cdots, \mathfrak{P}_k$ of $\mathfrak{p}$ in $\mathfrak{J}$. Here $\mathfrak{P}_\kappa$ can be defined as the set of those elements α of $\mathfrak{J}$ for which $F_\kappa(\alpha) = 0$. Two representations F_κ and F_λ belong to the same block if and only if $\mathfrak{P}_\kappa$ and $\mathfrak{P}_\lambda$ divide the same $\mathfrak{M}_\tau$.

11. Summary of the results. The principal aim of this paper is the proof of the following two theorems.

THEOREM 1. *Let $\mathfrak{G}$ be a group of order $g = p^a g'$, p a prime, $(g', p) = 1$. An ordinary irreducible representation Z_i of a degree $z_i \equiv 0 \pmod{p^a}$ remains irreducible as a modular representation, i.e. $\bar{Z}_i$ is equal to one of the F_κ, and $U_\kappa = F_\kappa$. Further Z_i forms a block B_τ of its own. The character $\zeta^{(i)}$ of Z_i vanishes for all elements of an order divisible by p.*

We denote a block B_τ of this kind as a block of highest kind. In the notations used above, we have here $x_\tau = y_\tau = 1$. We also shall show that for the blocks which are not of highest kind, we have $x_\tau > y_\tau$, in particular, $x_\tau > 1$.

THEOREM 2. *Let t_0 be the number of classes $\mathfrak{C}_\nu$ of conjugate elements in $\mathfrak{G}$ such that* (a) *the number of elements in $\mathfrak{C}_\nu$ is prime to p,* (b) *the elements of $\mathfrak{C}_\nu$ have an order prime to p. There exist exactly t_0 blocks B_τ which contain at least one ordinary irreducible representation Z_i of a degree z_i prime to p.*

We denote blocks of the type mentioned in this theorem as blocks of lowest kind. We also obtain some results for the blocks of intermediate types α, which contain only Z_i of degree $z_i \equiv 0 \pmod{p^\alpha}$ such that at least one of these degrees $z_i \not\equiv 0 \pmod{p^{\alpha+1}}$. The method which yields theorem 1 can, in a far more elaborate form, be used for a study of the blocks of type $a - 1$ as will be

[19] See R. A.

shown in another paper.[20] In the case $a = 1$, i.e. $g = p \cdot g'$, $(p, g') = 1$, each block is either of the highest or of lowest kind, so that the results give some information about every block. This can be made the basis for a study of this class of groups, which yields a large number of new results.[20] In order to attack the general group of finite order in a similar manner it would be necessary first to refine greatly the theory of blocks.

Two of the most important tools for the computation of the ordinary group characters are formed by the method of the multiplication of characters and the Frobenius' method of constructing characters of $\mathfrak{G}$ from characters of a subgroup $\mathfrak{H}$ of $\mathfrak{G}$. These methods can also be applied to modular characters. In part IV we study the former method, and in part V, the Frobenius' method. In part VI we consider a number of special cases and examples. We hope that in the results of these latter parts of our paper some justification can be seen for the somewhat complicated theory as developed in this lengthy introduction.

II. Blocks of Highest Kind

12. Condition for the reducibility of $\bar{Z}_i$. We use the same notations as in the introduction. If for one of the Z_i (§4) the corresponding modular representation $\bar{Z}_i$ becomes reducible, then there exists a non-singular matrix $\bar{M} = (\bar{m}_{ij})$ in the field $\bar{K}$ such that $\bar{M}^{-1}\bar{Z}_i\bar{M}$ breaks up into at least two constituents

$$\bar{M}^{-1}\bar{Z}_i\bar{M} = \begin{pmatrix} \bar{W}_1 & 0 \\ \bar{W}^3 & \bar{W}_4 \end{pmatrix}.$$

We choose a matrix $M = (m_{ij})$ such that m_{ij} lies in the residue class $\bar{m}_{ij}$ (mod $\mathfrak{p}$). Then the determinant of M is prime to $\mathfrak{p}$ and hence different from 0. Forming $M^{-1}Z_iM$, we obtain a formula

$$Z_i^* = M^{-1}Z_iM = \begin{pmatrix} W_1 & W_2 \\ W_3 & W_4 \end{pmatrix} \tag{36}$$

where all the coefficients in W_1, W_2, W_3, W_4, are $\mathfrak{p}$-integers of K, and those of W_2 are divisible by $\mathfrak{p}$ (i.e. each coefficient in W_2 is the quotient α/β of two integers of K such that $\alpha \equiv 0$, $\beta \not\equiv 0$ (mod $\mathfrak{p}$)). Let $Z_i^*(G) = (w_{\kappa\lambda}^{(i)}(G))$. According to the formulas of I. Schur,[21] we have[22]

$$\sum_G w_{\kappa\lambda}^{(i)}(G)w_{\rho\sigma}^{(i)}(G^{-1}) = g/z_i\delta_{\kappa\sigma}\delta_{\lambda\rho} \tag{37}$$

$$\sum_G w_{\kappa\lambda}^{(i)}(G)w_{\rho\sigma}^{(j)}(G^{-1}) = 0 \qquad \text{for } i \neq j. \tag{38}$$

In (37) we now take $\kappa = \sigma = 1$, $\lambda = \rho = z_i$ so that $\delta_{\kappa\sigma} = \delta_{\lambda\rho} = 1$. From the form of W_2 in (36), we have $w_{1z_i}(G) \equiv 0$ (mod $\mathfrak{p}$) for every G, hence $g/z_i \equiv 0$ (mod $\mathfrak{p}$) and consequently $g/z_i \equiv 0$ (mod p). Since g was exactly divisible by

[20] R. Brauer, Nat. Ac. of Sciences 25, 1939, p. 290.

[21] I. Schur, Sitzungsber. Preuss. Akad. 1905, p. 406.

[22] We set $\delta_{ij} = 0$ for $i \neq j$, $\delta_{ii} = 1$.

p^a (cf. (8)) we have $z_i \not\equiv 0 \pmod{p^a}$ if $\bar{Z}_i$ is reducible. This shows that if $z_i \equiv 0 \pmod{p^a}$, then $\bar{Z}_i$ is irreducible. This was the first part of theorem 1.

13. Blocks of highest kind. We now have to show that a Z_i with $z_i \equiv 0 \pmod{p^a}$ forms a block of its own. If this were not so, then we would have a Z_j with $i \neq j$ such that $\bar{Z}_i$ and $\bar{Z}_j$ have an irreducible constituent in common. But since $\bar{Z}_i$ itself is irreducible, $\bar{Z}_i$ would have to occur as constituent of $\bar{Z}_j$. Then there would exist a matrix $\bar{L} = (l_{ij})$ with coefficients in $\bar{K}$ such that[23]

$$\bar{L}^{-1}\bar{Z}_j\bar{L} = \begin{pmatrix} \bar{W}_1 & 0 & 0 \\ \bar{W}_4 & \bar{Z}_i & 0 \\ \bar{W}_7 & \bar{W}_8 & \bar{W}_9 \end{pmatrix}.$$

Choosing again the element l_{ij} in the residue class $l_{ij} \pmod{\mathfrak{p}}$, and setting $L = (l_{ij})$, we then have a formula

$$L^{-1}Z_jL = \begin{pmatrix} W_1 & W_2 & W_3 \\ W_4 & W_5 & W_6 \\ W_7 & W_8 & W_9 \end{pmatrix} \tag{39}$$

where

$$W_5 \equiv Z_i \pmod{\mathfrak{p}}. \tag{40}$$

We choose now $\kappa = \lambda = 1$ in (38), for ρ we take the number of the first row of W_5 in (39), for σ the number of the first column of W_5. Then (38) yields because of (40) and (37).

$$0 = \sum w_{11}^{(i)}(G)w_{\rho\sigma}^{(j)}(G^{-1}) \equiv \sum w_{11}^{(i)}(G)w_{11}^{(i)}(G^{-1}) = g/z_i \pmod{\mathfrak{p}}.$$

But this is impossible, because $z_i \equiv 0 \pmod{p^a}$. Consequently, no $\bar{Z}_j$ with $j \neq i$ belongs to the same block B_τ as Z_i. The block B_τ contains only the one ordinary irreducible representation Z_i, and only the one modular irreducible representation $\bar{Z}_i = F_\kappa$. In (28), C_τ and D_τ are matrices of degree 1, $x_\tau = y_\tau = 1$, and we have $D_\tau = 1$. Hence $C_\tau = 1$, because of (29). From (4), we obtain $U_\kappa = F_\kappa$.

14. Vanishing of the character for p-singular elements of $\mathfrak{G}$. Let H be a fixed element of $\mathfrak{G}$ the order of which is divisible by p. From the orthogonality relations for ordinary group characters, it follows that we have

$$\sum_{i=1}^{n} \zeta^{(i)}(G)\zeta^{(i)}(H) = 0$$

for every p-regular element G, since G and H^{-1} cannot be conjugate in $\mathfrak{G}$. Using (10), this can be written in the form

$$\sum_{i=1}^{n}\sum_{\kappa=1}^{k} d_{i\kappa}\varphi^{(\kappa)}(G)\zeta^{(i)}(H) = 0. \tag{41}$$

[23] The first row and column on the right side may be missing, also the last row and column.

Since (41) represents a linear relation between $\varphi^{(1)}(G), \varphi^{(2)}(G), \cdots, \varphi^{(k)}(G)$ which is true for every non-singular element G, the coefficient of each $\varphi^{(\kappa)}(G)$ must vanish

$$\sum_{i=1}^{n} d_{i\kappa}\zeta^{(i)}(H) = 0 \qquad (\kappa = 1, 2, \cdots, k).$$

If $\bar{Z}_i = F_\kappa$ as in §13, then $d_{i\kappa} = 1$, and $d_{j\kappa} = 0$ for $i \neq j$ since F_κ does not appear in $\bar{Z}_j$. Hence $\zeta^{(i)}(H) = 0$. This proves theorem 1 completely.

In the case $(g, p) = 1$, it follows at once from theorem 1 that each ordinary irreducible representation Z_i remains irreducible when taken as a modular representation and so $\bar{Z}_i$ is equivalent to some F_κ. Then $k = n$ and the $\zeta^{(i)}$ are identical with the $\varphi^{(\kappa)}, \eta^{(\kappa)}$. The relations (16) and (20) here become the same as the Frobenius relations for group characters (cf. §§3, 8.)

15. Example. As an example, we mention the simple Mathieu group M_{12} of order $12.11.10.9.8 = 2^6.3^3.5.11 = 95040$, the characters of which have been given by Frobenius.[24] The degree of the ordinary irreducible representations are

1, 11, 11, 16, 16, 45, 54, 55, 55, 55, 66, 99, 120, 144, 176.

Of these 15 characters, 8 are of highest kind (mod 11), for instance, the character of degree 176. From the table of characters, it follows that this character is the product of a character of degree 11 and a character of degree 16. Consequently the characters of degree 16 must also remain irreducible (mod 11), since a splitting would imply a splitting of the character of degree 176. The characters of degree 16 are not of highest kind.

For $p = 5$ we have 5 characters of highest kind, for $p = 3$ there is one of them and for $p = 2$ there is no such character.

III. The Elementary Divisors of C

16. Computation of the elementary divisors of C. In the following section we work in the ring $\mathfrak{o}_\mathfrak{p}$ of $\mathfrak{p}$-integers of K. If π is an element such that $\pi \equiv 0 \pmod{\mathfrak{p}}$, $\pi \not\equiv 0 \pmod{\mathfrak{p}^2}$, then every ideal of $\mathfrak{o}_\mathfrak{p}$ is of the form $(\pi)^m$, and therefore, the theory of elementary divisors holds for matrices with coefficients in $\mathfrak{o}_\mathfrak{p}$. In formula (15), $\Phi' C \Phi = \mathrm{T}$, the determinant of Φ is a unit of $\mathfrak{o}_\mathfrak{p}$ because of (17). Consequently, C and T have the same elementary divisors. But the elementary divisors of T can be obtained directly from (14), they are the highest powers of p which divide the numbers

$$g/g_1, g/g_2, \cdots, g/g_k. \tag{42}$$

We now consider C as a matrix with coefficients in the ring of rational integers, and denote the elementary divisors corresponding to this case by $e_1, e_2, \cdots, e_k$. It follows that *the powers of p which divide these integers are exactly the same*

[24] The ordinary characters of the Mathieu-groups have been given by G. Frobenius, Sitzungsber. Preuss. Akad. 1904, p. 558.

powers which appear in the integers g/g_ν, $\nu = 1, 2, \cdots, k$, if the latter are properly arranged.

In another paper, it will be shown that the determinant of C is actually a power of p. Then it follows that the e_ν are themselves the powers of p which divide the numbers (42). For our present purpose this finer result will not be needed.

17. Blocks of type α. We say that a block B_τ is of type α, if it contains only representations F_κ of degrees $f_\kappa \equiv 0 \pmod{p^\alpha}$, and if at least one of these degrees is not divisible by $p^{\alpha+1}$.

By (12)

$$u_\kappa = \sum_\lambda c_{\kappa\lambda} f_\lambda \qquad (\lambda = 1, 2, \cdots, k).$$

Because of the form (28) of C, the corresponding formulas hold, when we restrict κ and λ to those values for which F_κ, F_λ belong to the block B_τ.

We can find two unimodular matrices M_1 and M_2 such that $C_\tau^* = M_1 C_\tau M_2$ has zeros outside of the main diagonal, and contains the elementary divisors e_κ in the main diagonal.[25] We have then

$$u_\kappa^* = e_\kappa f_\kappa^* \tag{43}$$

where the u_κ^* are obtained from the u_λ by the linear transformation M_1, and the f_κ^* from the f_λ by the transformation M_2^{-1}; κ, λ range over the values corresponding to B_τ. Since M_2 is unimodular, all the f_κ^* are divisible by p^α, and one of them is not divisible by $p^{\alpha+1}$. On the other hand, the u_κ^* are divisible by p^a, according to (18). From (43) it follows that *at least one of the elementary divisors e_κ corresponding to B_τ is divisible by $p^{a-\alpha}$.* Let s_α denote the number of blocks of type α, and a_α the number of integers g_ν ($\nu = 1, 2, \cdots, k$) which are divisible by p^α and not by $p^{\alpha+1}$. We consider now the $s_\alpha + s_{\alpha-1} + \cdots + s_0$ blocks of type $\leqq \alpha$. To each of them corresponds an elementary divisor which is divisible at least by $p^{a-\alpha}$ and hence a number g/g_ν which is at least divisible by this number. Then g_ν will be divisible at most by p^α, and we find

$$s_0 + s_1 + \cdots + s_\alpha \leqq a_0 + a_1 + \cdots + a_\alpha .$$

THEOREM 3. *Let $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_k$ be the classes of conjugate p-regular elements in $\mathfrak{G}$ and denote by g_ν the number of elements in $\mathfrak{C}_\nu$. If a_α of the numbers g_ν are divisible by p^α and not by $p^{\alpha+1}$, and if $\mathfrak{G}$ possesses s_α blocks of type α, then* (for $\alpha = 0, 1, 2, \cdots, a$)

$$s_0 + s_1 \cdots + s_\alpha \leqq a_0 + a_1 \cdots + a_\alpha . \tag{44}$$

18. Ordinary characters which are linearly independent mod $\mathfrak{p}$.

LEMMA: There exist k but not more than k ordinary irreducible characters $\zeta^{(i)}$ which are linearly independent (mod $\mathfrak{p}$).

[25] The elementary divisors of all the C_τ together are $e_1, e_2, \cdots, e_k$ (in some arrangement) because of (28).

Proof: If a linear relation

$$\sum_i a_i \zeta^{(i)} \equiv 0 \pmod{\mathfrak{p}}$$

with p-integral coefficients holds for the p-regular classes, then it holds for every class. This can be seen similarly as in §6. From this remark, it follows already that we cannot have more than k ordinary characters which are linearly independent (mod $\mathfrak{p}$).

For the proof that there exist k such independent characters, we use a method by which one of us showed earlier the existence of k irreducible modular characters.[26] If the maximal number of ordinary irreducible characters, which are linearly independent mod $\mathfrak{p}$, was smaller than k, then we could find $\mathfrak{p}$-integers b_ν such that

$$\sum_{\nu=1}^{k} b_\nu \zeta_\nu^{(i)} \equiv 0 \pmod{\mathfrak{p}}$$

and the b_ν are not all divisible by $\mathfrak{p}$. Since a reducible character of $\mathfrak{G}$ is a sum of irreducible character $\zeta^{(i)}$, we would have

$$\sum_{\nu=1}^{k} b_\nu \zeta_\nu \equiv 0 \pmod{\mathfrak{p}} \tag{45}$$

for any character ζ of $\mathfrak{G}$. We want to show that this is impossible if the b_ν are not all divisible by $\mathfrak{p}$. We assume that the corresponding result for all proper subgroups $\mathfrak{H}$ of $\mathfrak{G}$ has already been shown.

Let $\mathfrak{H}$ be a proper subgroup of $\mathfrak{G}$, let $\mathfrak{C}_1', \mathfrak{C}_2', \cdots, \mathfrak{C}_l'$ be the classes of conjugate p-regular elements of $\mathfrak{H}$ and let H_γ be an element of $\mathfrak{C}_\gamma'$. If ψ is any character of $\mathfrak{H}$, we set $\psi(G) = 0$ if G does not belong to $\mathfrak{H}$. We determine a (right hand side) residue system $P_1, P_2, \cdots, P_m$ of $\mathfrak{G}$ mod $\mathfrak{H}$. According to Frobenius.[27]

$$\zeta(G) = \sum_{\mu=1}^{m} \psi(P_\mu G P_\mu^{-1}) \tag{46}$$

is a character of $\mathfrak{G}$. Since (45) hold for every character of $\mathfrak{G}$, it holds for (46). The elements G and $P_\mu G P_\mu^{-1}$ are p-regular at the same time. If G_ρ is an element of $\mathfrak{C}_\rho$, we have from (46)

$$\zeta_\rho = \sum_{\sigma=1}^{l} l_{\rho\sigma} \psi_\sigma \tag{47}$$

where $l_{\rho\sigma}$ denotes the number of P_μ for which $P_\mu G_\rho P_\mu^{-1}$ is conjugate to H_σ with regard to $\mathfrak{H}$. For each H_σ there exists exactly one ρ for which $l_{\rho\sigma} \neq 0$ since one class $\mathfrak{C}_\rho$ in $\mathfrak{G}$ must contain H_σ. We denote this ρ by $\tau(\sigma)$. From (45) and (47) it follows that

$$\sum_{\sigma=1}^{l} \left(\sum_{\rho=1}^{k} b_\rho l_{\rho\sigma} \right) \psi_\sigma \equiv 0 \pmod{\mathfrak{p}}. \tag{48}$$

[26] See footnote 15.

[27] G. Frobenius, Sitzungsber. Preuss. Akad. 1898, p. 501.

This must hold for every character ψ of $\mathfrak{H}$. But (48) represents a congruence of exactly the same type for $\mathfrak{H}$, as (45) has for $\mathfrak{G}$. According to our assumption concerning $\mathfrak{H}$, the coefficients of every ψ_σ must be divisible by $\mathfrak{p}$,

$$\sum_{\rho=1}^{k} b_\rho l_{\rho\sigma} \equiv 0 \pmod{\mathfrak{p}} \qquad \text{for } \sigma = 1, 2, \cdots, l$$

and since only $l_{\tau(\sigma),\sigma} \not\equiv 0$

$$b_{\tau(\sigma)} l_{\tau(\sigma),\sigma} \equiv 0 \pmod{\mathfrak{p}} \qquad \text{for } \sigma = 1, 2, \cdots, l. \tag{49}$$

So far the subgroup $\mathfrak{H}$ has been arbitrary. We now try to determine $\mathfrak{H}$ for a given value ρ, $(\rho = 1, 2, \cdots, k)$, such that (a) ρ appears in the form $\rho = \tau(\sigma)$, and (b) for this σ the number $l_{\rho\sigma}$ is not divisible by $\mathfrak{p}$. Then (49) implies $b_\rho \equiv 0 \pmod{\mathfrak{p}}$, and if this holds for every ρ, then we have arrived at a contradiction with the fact that the congruence (45) was to be not trivial.

The condition (a) is satisfied when G_ρ belongs to $\mathfrak{H}$, we may take $G_\rho = H_\sigma$. If we choose $\mathfrak{H}$ as a subgroup of the normalizer $\mathfrak{N}$ of G_ρ, then G_ρ is only conjugate to itself with regard to $\mathfrak{H}$. Then $l_{\rho\sigma}$ in (46) can be defined as the number of P_μ for which $P_\mu G_\rho P_\mu^{-1} = G_\rho$. If $\mathfrak{H}$ has the order h, then $hl_{\rho\sigma} = N$ is the order of $\mathfrak{N}$. We have only to take care that $\mathfrak{H}$ contains a p-Sylow group of $\mathfrak{N}$. Then h is divisible by the same power of p as N, hence $l_{\rho\sigma} \not\equiv 0 \pmod{\mathfrak{p}}$ and, therefore, condition (b) is satisfied.

We can, therefore, satisfy the above conditions (a) and (b) by choosing $\mathfrak{H}$ as the subgroup which is generated by G_ρ and a p-Sylow group of the normalizer $\mathfrak{N}$ of G_ρ. Here, however, an exceptional case is possible which must be treated separately. The group defined in this manner can be identical with $\mathfrak{G}$.

In this case, the only p-regular elements of $\mathfrak{G}$ are $1, G_\rho, G_\rho, \cdots, G_\rho^{q-1}$, where q is the order of G_ρ. We obtain a character of $\mathfrak{G}$ by associating ϵ^μ with G^μ where ϵ is a q-th root of unity. Then (45) becomes

$$\sum_{\mu}^{q-1} b_\mu \epsilon^\mu \equiv 0 \pmod{\mathfrak{p}}.$$

We multiply here with $\epsilon^{-\beta}$ for a fixed β and add over all q-th roots of unity. Since $(q, p) = 1$ we find $b_\beta \equiv 0 \pmod{\mathfrak{p}}$ for $\beta = 0, 1, 2, \cdots, q - 1$, which gives a contradiction.

19. Applications of the lemma. It follows immediately from the lemma in §18 that the congruences

$$\sum_{i=1}^{n} a_i \zeta_\nu^{(i)} \equiv \eta_\nu \pmod{\mathfrak{p}} \qquad (\text{for } \nu = 1, 2, \cdots, k) \tag{50}$$

can be solved with regard to $a_1, a_2, \cdots, a_n$ if $\eta_1, \eta_2, \cdots, \eta_k$ are any given $\mathfrak{p}$-integers of K. The a_i also will be $\mathfrak{p}$-integers of K.

From (10) and (28) the number of $\zeta^{(i)}$ in a given block B_τ which are linearly independent mod $\mathfrak{p}$, is at most equal to the number y_τ of modular characters

$\varphi^{(\kappa)}$ in B_τ. But since $y_1 + \cdots + y_s$ is the full number k of modular irreducible characters, this implies that B_τ contains y_τ characters $\zeta^{(i)}$ which are linearly independent (mod $\mathfrak{p}$). It follows that the matrix D_τ of type (x_τ, y_τ) in (28) still has the rank y_τ when it is considered mod $\mathfrak{p}$.

From this remark and (10) it follows that the modular characters $\varphi^{(\kappa)}$ can be expressed by means of the ordinary characters with p-integral rational coefficients. For a block of type α, all the f_κ are divisible by p^α. Since by (12), $z_i = \sum_\kappa d_{i\kappa} f_\kappa$, it follows that all the z_i of the block B_τ will be divisible by p^α. On the other hand, the z_i of B_τ cannot all be divisible by $p^{\alpha+1}$, since otherwise all the f_κ of B_τ would be divisible by $p^{\alpha+1}$, as we see when we express the $\varphi^{(\kappa)}$ of B_τ as linear combinations of the $\zeta^{(i)}(G)$ of B_τ with p-integral coefficients and set $G = 1$. *We can define a block B_τ of type α by the fact that the degrees of the ordinary irreducible characters of B_τ are all divisible by p^α but not by $p^{\alpha+1}$.* In the definition in §17 we can replace the modular characters by ordinary characters. In particular, the blocks of type 0 are the blocks of lowest kind; the blocks of type a, the blocks of highest kind (§11).

If the $\zeta^{(i)}$ of B_τ are arranged in a suitable order, then the first y_τ of them will be linearly independent mod $\mathfrak{p}$. We may then find a matrix V of degree y_τ with p-integral rational coefficients and a determinant prime to p such that

$$D_\tau V = \begin{pmatrix} I \\ M \end{pmatrix}$$

where M is a matrix of type $(x_\tau - y_\tau, y_\tau)$ and I, the unit matrix of degree y_τ. Using (29), we find

$$V'C_\tau V = V'D_\tau' D_\tau V = (I, M')\begin{pmatrix} I \\ M \end{pmatrix} = I + M'M. \tag{51}$$

We work in the Galois field with p elements, replacing every number by its residue class (mod p). If M in this sense has rank m, then we can find $y_\tau - m$ linearly independent vectors ξ of y_τ dimensions for which $M\xi = 0$. For these vectors, we have $(I + M'M)\xi = \xi$ so that at least $y_\tau - m$ linearly independent vectors are obtained in the form $(I + M'M)\eta$ where η is an arbitrary vector. It follows that $(I + M'M)$ has (mod p) a rank $r \geqq y_\tau - m$. Because of (51) C_τ has (mod p) the same rank r. Then exactly r of the elementary divisors of C_τ will be not divisible by p. But $m \leqq x_\tau - y_\tau$, since M has $x_\tau - y_\tau$ rows, so

$$r \geqq y_\tau - (x_\tau - y_\tau) = 2y_\tau - x_\tau.$$

THEOREM 4. *If the block B_τ contains x_τ ordinary and y_τ modular irreducible characters, then the corresponding part C_τ of the Cartan matrix has at least $2y_\tau - x_\tau$ elementary divisors which are not divisible by p (and hence equal to* 1 *according to the theorem quoted in* §16).

If $2y_\tau < x_\tau$ then this theorem does not give anything.

If for a block we have $y_\tau = x_\tau$, then x_τ of the elementary divisors will be prime to p. But this is the total number of elementary divisors of B_τ. Ac-

cording to §17, this is impossible if B_τ is of type $\alpha < a$. Hence B_τ is of highest type, and then $x_\tau = y_\tau = 1$ (theorem 1). Using (30) we obtain

THEOREM 5: *Every block, which is not of highest kind, contains more ordinary than modular irreducible characters.*[28]

This shows that Theorem 1 characterizes the blocks of highest kind.

20. Blocks of lowest kind. We now come to the proof of theorem 2 (§11). Let $\eta_1, \eta_2, \cdots, \eta_k$ be any given $\mathfrak{p}$-integers of K. We solve the congruences (50). Let $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_m$ be these classes of p-regular elements of $\mathfrak{G}$, for which the number g_ν of elements in the class is not divisible by p, then

$$m = a_0 \tag{52}$$

where a_0 was defined in §17.

The number $\omega_\nu^{(i)} = g_\nu \zeta_\nu^{(i)}/z_i$ (cf. (24) and (26)) is an algebraic integer. If $z_i \equiv 0$ (mod p) and $(g_\nu, p) = 1$, then $\zeta_\nu^{(i)} \equiv 0 \pmod{\mathfrak{p}}$. The corresponding terms in (50) can be omitted. We have, therefore,

$$\sum_i a_i \zeta_\nu^{(i)} \equiv \eta_\nu \pmod{\mathfrak{p}} \qquad \text{for } \nu = 1, 2, \cdots, m \tag{53}$$

where the sum is extended over such values of i, for which $z_i \not\equiv 0$ (mod p). In particular, only characters of blocks of the lowest kind appear. We can pick out one character $\zeta^{(h)}$ in B_τ such that $z_h \not\equiv 0$ (mod p). If $\zeta^{(i)}$ is another character of B_τ, then according to (26) and (27) we have

$$\frac{g_\nu \zeta_\nu^{(i)}}{z_i} \equiv \frac{g_\nu \zeta_\nu^{(h)}}{z_h} \pmod{\mathfrak{p}}$$

$$\zeta_\nu^{(i)} \equiv \frac{z_i}{z_h} \zeta_\nu^{(h)}, \pmod{\mathfrak{p}}, \qquad (\nu = 1, 2, \cdots, m)$$

since $(g_\nu, p) = 1$, for $\nu = 1, 2, \cdots, m$. We substitute this value in (53) and obtain formulae

$$\sum_h b_h \zeta_\nu^{(h)} \equiv \eta_\nu \pmod{\mathfrak{p}}, \qquad (\nu = 1, 2, \cdots, m) \tag{54}$$

where the $\zeta^{(h)}$ are the characters we selected in the blocks of lowest kind. The number of terms on the left side then is the number s_0 of blocks of the lowest kind. The b_h are $\mathfrak{p}$-integers which are independent of ν. For every given set of $\mathfrak{p}$-integers $\eta_1, \eta_2, \cdots, \eta_m$ the congruences have a solution b_h. The number of unknowns b_h cannot be smaller than the number of congruences, hence, from (52)

$$a_0 = m \leqq s_0 .$$

But from (44), it follows that $s_0 \leqq a_0$. Consequently, $s_0 = a_0$, and this is exactly the statement of theorem 2, §11.

[28] A second proof for this relation $x_\tau > y_\tau$ is given in §27. It also can be proved by considering the representation of the elements of the center of the group ring.

In other words this result can be expressed as follows: An elementary divisor divisible by p^a can appear in C_τ only if the block B_τ is of lowest kind. In this case there is exactly one such elementary divisor.

We easily see now that for blocks of type 0, 1, and a there is at least one U_κ of the block whose degree $u_\kappa \not\equiv 0 \pmod{p^{a+1}}$. For a block B_τ of type α it follows from (43) that if all $u_\kappa \equiv 0 \pmod{p^{a+1}}$ then at least one elementary divisor of C_τ is divisible by $p^{a-\alpha+1}$. For $\alpha = 0$ this would mean an elementary divisor divisible by p^{a+1} which is not possible (cf. §16). For $\alpha = 1$ it means an elementary divisor divisible by p^a but by the above statement of theorem 2 such divisors can appear only for blocks of type 0. In case $\alpha = a$ the remark is obvious, since the block is then of highest kind. For values of α intermediate to 1 and a we can only as yet say that an elementary divisor of C_τ is divisible at most by p^{a-1}, and hence from (43) at least one u_k is divisible at most by $p^{a+\beta}$ where $\beta \leqq \alpha - 1$.

21. Alternative proof of theorem 2. We here begin with the components $(\tilde{\gamma}_{\kappa\lambda})$ of the matrix gC^{-1} (cf. §9), and again work in the ring of all rational p-integers. If the matrix C_τ (cf. (28)) has the elementary divisors p^{α_ν}, $(\nu = 1, 2, \cdots, y_\tau)$, then gC_τ^{-1} has the elementary divisors $p^{a-\alpha_\nu}$. In the case that B_τ is not a block of the lowest kind, then for any pair F_κ, F_λ belonging to B_τ the degrees f_κ, f_λ are divisible by p, and so from (34) $\tilde{\gamma}_{\kappa\lambda} \equiv 0 \pmod p$. Hence all the α_ν are smaller than a in this case. If B_τ is of the lowest kind, then since by (34) the matrix $gC_\tau^{-1} \equiv (f_\kappa f_\lambda S_\tau)$, ($\kappa$ row index, λ column index) the rank of gC_τ^{-1} (mod p) is 1 or 0 according as to whether $S_\tau \not\equiv 0 \pmod p$, or $S_\tau \equiv 0 \pmod p$. The considerations in §17 show that for a block of the lowest kind, at least one of the elementary divisors of C_τ is $\geqq p^a$, $\alpha_\nu \geqq a$. It follows that gC_τ^{-1} has one elementary divisor 1, and we have

$$S_\tau \not\equiv 0 \pmod p \quad (B_\tau \text{ block of the lowest kind}). \tag{55}$$

Since here gC_τ^{-1} has (mod p) one elementary divisor 1, C_τ has exactly one elementary divisor p^a. Consequently, the number of blocks of the lowest kind is equal to the number of elementary divisors p^a of C. This, in connection with the result of §16 yields theorem 1.

We add some remarks about the determination of the numbers S_τ for blocks of the lowest kind. From (12) it follows that $f_\kappa = g^{-1} \sum_\lambda \tilde{\gamma}_{\kappa\lambda} u_\lambda$. Combining this with (34), we obtain

$$f_\kappa = \frac{1}{g'} \sum \tilde{\gamma}_{\kappa\lambda} \frac{u_\lambda}{p^a} \equiv \frac{1}{g'} f_\kappa S_\tau \sum{}' f_\lambda \frac{u_\lambda}{p^a} \pmod p$$

where λ ranges over all values for which $\varphi^{(\lambda)}$ belongs to the block B_τ. If B_τ is of the lowest kind, we may assume $f_\kappa \not\equiv 0 \pmod p$. Hence

$$g' \equiv S_\tau \sum{}' \frac{f_\lambda u_\lambda}{p^a} \pmod p \qquad (\varphi^{(\lambda)} \text{ in } B_\tau) \tag{56}$$

whence S_τ (mod p) can be obtained, if only the degrees of the characters are known. Using (12) and (5), we easily obtain

$$\sum_i{}'' z_i^2 = \sum_i{}'' \sum_\lambda{}' (d_{i\lambda} f_\lambda)^2 = \sum_{\kappa,\lambda}{}' c_{\kappa\lambda} f_\kappa f_\lambda = \sum_\kappa{}' u_\kappa f_\kappa \tag{57}$$

where i ranges over those values for which $\zeta^{(i)}$ belongs to B_τ. Hence

$$g' \equiv S_\tau \frac{\sum'' z_i^2}{p^a} \pmod{p}. \tag{58}$$

The numbers S_τ can also be determined in a different manner from the ordinary group characters $\zeta^{(i)}$ of $\mathfrak{G}$. We set

$$V = (v_{ij}) = \left(\sum_{\nu=1}^{k} g_\nu \zeta_\nu^{(i)} \zeta_{\nu^*}^{(j)}\right). \tag{59}$$

We have then, making use of (21)

$$v_{ij} = \sum_{\nu=1}^{k} \sum_{\kappa=1}^{k} \sum_{\lambda=1}^{k} g_\nu d_{i\kappa} d_{j\lambda} \varphi_\nu^{(\kappa)} \varphi_{\nu^*}^{(\lambda)} = \sum_{\kappa,\lambda=1}^{k} d_{i\kappa} d_{j\lambda} \tilde{\gamma}_{\kappa\lambda}$$

$$V = gDC^{-1}D' = gD(D'D)^{-1}D'. \tag{60}$$

Using (34), we obtain

$$v_{ij} \equiv \sum_{\kappa,\lambda} d_{i\kappa} d_{j\lambda} f_\kappa f_\lambda S_\tau \pmod{p}$$

if $\zeta^{(i)}$ and $\zeta^{(j)}$ both belong to B_τ. But $\sum_\kappa d_{i\kappa} f_\kappa = z_i$, according to (12), and hence

$$\begin{aligned} v_{ij} &\equiv z_i z_j S_\tau \pmod{p} && (\zeta^{(i)} \text{ and } \zeta^{(j)} \text{ in } B_\tau) \\ v_{ij} &= 0 && (\zeta^{(i)} \text{ and } \zeta^{(j)} \text{ in different blocks}). \end{aligned} \tag{61}$$

IV. On the Multiplication of the Characters

22. Relations between the problems of determining the ordinary and the modular characters of $\mathfrak{G}$. For any group $\mathfrak{G}$ of order g, we have the two problems of finding the ordinary irreducible characters $\zeta^{(i)} (i = 1, 2, \cdots, n)$ and the modular irreducible characters $\varphi^{(\kappa)}$ $(k = 1, 2, \cdots, k)$ for a fixed prime p. We may assume that p divides g, since otherwise the two types of characters coincide. We ask now: (a) How much does knowing the ordinary characters help in the determination of the modular characters? (b) How much does knowing the modular characters help in the determination of the ordinary characters? It seems that in general we obtain some valuable information, but that in neither case the complete answer can be found. For instance, in the case of a p-group, the modular characters become trivial, since there is only the (1)-character, and this shows clearly that we cannot expect that the $\zeta_\nu^{(i)}$ are determined uniquely by the $\varphi_\nu^{(\kappa)}$. For both questions (a) and (b), it is of course of great importance to find the matrix D (cf. (13)).

If the $\varphi_\nu^{(\kappa)}$ are known, then (21) permits the determination of the matrix C, so that we also may find the characters $\eta^{(\kappa)}$ of the indecomposable constituents U_κ. For the determination of D, we have the formulas (5). In certain cases, these formulas are sufficient to find D, cf. the example of the group LF (2, p) in §31. But, in general, we must expect several possible solutions for D some of which may belong to other groups $H, K, \cdots$ which also have $(\varphi_\lambda^{(\kappa)})$ as their modular characters. There is, of course, only a finite number of possibilities for D. If D itself is known, then the values of the ordinary characters $\zeta_\nu^{(i)}$ for p-regular elements G of $\mathfrak{G}$ can be obtained from (10). There remains then the determination of the values of the characters for the other classes. Mod p, we can find these values from the values of the characters for the p-regular classes (cf. §§6, 18). Further, we obtain conditions from the orthogonality relations for group characters. Also the method of multiplying characters can be used with advantage. It may be mentioned that in many important cases it seems easier to find the modular characters than the ordinary characters. For instance, in the case of many simple groups, the analogy with semisimple continuous groups can be used in the modular theory.

Conversely, let us assume now that the ordinary characters $\zeta^{(i)}$ are known. It follows from (13) that $D = Z\Phi^{-1}$, which shows that each column of D is of the form

$$d_i = \sum_{\nu=1}^{k} \zeta_\nu^{(i)} \alpha_\nu, \tag{62}$$

that is, each column of D is a linear combination of those columns in the tables of ordinary characters which correspond to p-regular elements. The α_ν are the elements of a column of Φ^{-1} and, therefore, are not known, but we have some information about them. For instance, they are of form β/g' where β is an integer of the field generated by the $\zeta^{(i)}$ and $g = p^a g'$, $(g', p) = 1$. The d_i must be rational integers $\geqq 0$. Further restrictions are obtained from (13) and the form (28) of D, and from the fact that the determinant of $D'D$ is known. But these conditions are not enough to determine D uniquely, several cases will have to be considered. If D is known, then the modular characters are known. The equations (59) and (60) show that the matrix $D(D'D)^{-1}D'$ can be found if the $\zeta^{(i)}$ are known, but this does not provide any new information.

We may add some remarks in this connection. The condition (62) is, of course, equivalent to saying that each column of D is orthogonal to each column of Z which corresponds to a p-singular element of G. If a vector $x = (x_1, x_2, \cdots, x_n)$ is orthogonal to all these columns of Z, then x is a linear combination of the columns of D. Similarly, if $y = (y_1, y_2, \cdots, y_n)$ is orthogonal to all columns of D, then y is a linear combination of the columns of Z which correspond to p-singular elements of $\mathfrak{G}$.

If a relation

$$\sum_{i=1}^{n} \zeta^{(i)}(G)\beta_i = 0 \quad \text{(for all } p\text{-regular elements } G \text{ of } \mathfrak{G}) \tag{63}$$

where the β_i are independent of G, then $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$ is a linear combination of the columns of Z which correspond to p-singular elements, and vice versa. Hence β is orthogonal to every column of D. We cut $(\beta_1, \beta_2, \cdots, \beta_n)$ into s pieces corresponding to the s blocks B_τ, and replace all the β_i by 0 except those which belong to a fixed piece. This modified vector β still is orthogonal to all the columns of D because of the form (28) of D. Hence the modified vector β still satisfies (63)

$$\sum{}' \zeta^{(i)}(G)\beta_i = 0$$

where i ranges over only those values for which $\zeta^{(i)}$ belongs to a fixed block B_τ.

THEOREM 6. *If a linear relation between the ordinary characters holds for all p-regular elements of $\mathfrak{G}$, then the relation remains true if we leave away all terms except those which contain the characters of a fixed block B_τ.*

23. The multiplication of characters. If F and H are two modular representations then $G \to F(G) \times H(G)$ gives a new representation $F \times H$. Since the characteristic roots of the Kronecker product $F(G) \times H(G)$ are obtained by multiplying each characteristic root of $F(G)$ into each characteristic root of $H(G)$, it follows easily that the character of $F \times H$ is obtained by multiplying the characters of F and H. Applying this to the irreducible characters $\varphi^{(\kappa)}$, $\varphi^{(\lambda)}$ we have that $\varphi^{(\kappa)} \cdot \varphi^{(\lambda)}$ is again a character of $\mathfrak{G}$, reducible or irreducible, and we obtain formulas

$$\varphi^{(\kappa)} \cdot \varphi^{(\lambda)} = \sum_\mu a_{\kappa\lambda\mu} \varphi^{(\mu)} \tag{64}$$

where the $a_{\kappa\lambda\mu}$ are rational integers, $a_{\kappa\lambda\mu} \geqq 0$. There is, of course some connection with the corresponding coefficients appearing in the multiplication of ordinary characters. If we have

$$\zeta^{(i)}\zeta^{(j)} = \sum_k b_{ijk}\zeta^{(k)} \tag{65}$$

then we express ζ by means of the φ (cf. (10)) and obtain

$$\zeta^{(i)}\zeta^{(j)} = \sum d_{i\kappa} d_{j\lambda} \varphi^{(\kappa)} \varphi^{(\lambda)} = \sum d_{i\kappa} d_{j\lambda} a_{\kappa\lambda\mu} \varphi^{(\mu)}$$

$$\zeta^{(i)}\zeta^{(j)} = \sum b_{ijk}\zeta^{(k)} = \sum b_{ijk} d_{k\mu} \varphi^{(\mu)}$$

$$\sum_{\kappa,\lambda} d_{i\kappa} d_{j\lambda} a_{\kappa\lambda\mu} = \sum_k b_{ijk} d_{k\mu}. \tag{66}$$

We derive some further relations for the $a_{\kappa\lambda\mu}$. Of course, they must satisfy the conditions for the constants of multiplication of a commutative algebra

$$a_{\kappa\lambda\mu} = a_{\lambda\kappa\mu}, \qquad \sum_\mu a_{\kappa\lambda\mu} a_{\mu\tau\sigma} = \sum_\mu a_{\kappa\mu\sigma} a_{\lambda\tau\mu}.$$

To each representation of $\mathfrak{G}$ there corresponds a contragredient representation. We denote the representation contragredient to F_κ by $F_{\kappa'}$. Then

$$\varphi_\nu^{(\kappa')} = \varphi_{\nu^*}^{(\kappa)} \tag{67}$$

where $\mathfrak{C}_{\nu^*}$ as in §7 denotes the class reciprocal to $\mathfrak{C}_\nu$. Here $1', 2', \cdots, k'$ and $1^*, 2^*, \cdots, k^*$ are permutations of period 2 of $1, 2, \cdots, k$. From (64) we obtain

$$a_{\kappa'\lambda'\mu'} = a_{\kappa\lambda\mu} \tag{68}$$

since the contragredient of $\varphi^{(\kappa)} \cdot \varphi^{(\lambda)}$ is $\varphi^{(\kappa')} \cdot \varphi^{(\lambda')}$. Further, the regular representation R is self-contragredient. Hence the contragredient of U_κ is an indecomposable constituent of R, and since its irreducible top constituent is $F_{\kappa'}$, we see that $U_{\kappa'}$ and U_κ are contragredient. This implies

$$c_{\kappa\lambda} = c_{\kappa'\lambda'}. \tag{69}$$

From (64) and the orthogonality relations (20) we obtain

$$g a_{\kappa\lambda\mu} = \sum_{\nu=1}^{k} g_\nu \varphi_\nu^{(\kappa)} \varphi_\nu^{(\lambda)} \eta_{\nu^*}^{(\mu)}. \tag{70}$$

By multiplying the two left hand members through by $\gamma_{\mu\rho'} = \gamma_{\rho'\mu}$ and adding over μ we obtain

$$\sum_\mu g a_{\kappa\lambda\mu} \gamma_{\mu\rho'} = \sum_{\nu=1}^{k} g_\nu \varphi_\nu^{(\kappa)} \varphi_\nu^{(\lambda)} \varphi_{\nu^*}^{(\rho')}$$

or

$$\sum_\mu a_{\kappa\lambda\mu} \tilde{\gamma}_{\mu\rho'} = \sum_{\nu=1}^{k} g_\nu \varphi_\nu^{(\kappa)} \varphi_\nu^{(\lambda)} \varphi_\nu^{(\rho)}$$

which shows that the left side remains unchanged, when the three indices κ, λ, ρ are permuted. Thus

$$\sum a_{\kappa\lambda\mu} \tilde{\gamma}_{\mu\rho'} = \sum a_{\rho\lambda\mu} \tilde{\gamma}_{\mu\kappa'}. \tag{71}$$

In particular, for $\kappa = 1$, we have $a_{\kappa\lambda\mu} = \delta_{\lambda\mu}$, $\kappa' = 1$, and when we interchange ρ and ρ' we have

$$\tilde{\gamma}_{\lambda\rho} = \sum_\mu a_{\rho'\lambda\mu} \tilde{\gamma}_{\mu 1} \tag{72}$$

which shows that the whole matrix C^{-1} can be found if its first column and the constants of multiplication are known.

If $\varphi^{(\lambda)} \varphi^{(\rho')}$ does not contain a character of the block B_1, then the right side of (72) vanishes and $\tilde{\gamma}_{\lambda\rho} = 0$.

THEOREM 7.[29] *If the product of a character $\varphi^{(\lambda)}$ with the contragredient character $\varphi^{(\rho')}$ of $\varphi^{(\rho)}$ does not contain a character of the first block B_1, then the corresponding coefficient $\tilde{\gamma}_{\lambda\rho}$ of gC^{-1} vanishes.*

If the block B_τ of $\varphi^{(\lambda)}$ contains more than one modular character then because of the form (28) of C, $\tilde{\gamma}_{\lambda\rho} = 0$, cannot hold for all $\rho \neq \lambda$ such that $\varphi^{(\rho)}$ belongs

[29] This theorem is related to theorem 2 of R. Brauer, Math. Zeitschr. 41, 1936, p. 330.

to B_τ. Further, if $\varphi^{(\lambda)}$ and $\varphi^{(\rho)}$ belong to the same block, and if their degrees are not divisible by p, then $\tilde{\gamma}_{\lambda\rho} \neq 0$ according to (34) and (55). Hence we have the

COROLLARY. *If two characters $\varphi^{(\lambda)}$ and $\varphi^{(\rho)}$ belong to the same block and both have degrees prime to p, then $\varphi^{(\lambda)} \cdot \varphi^{(\rho)}$ contains a character of the first block.*

We prove two more formulas connecting the $c_{\kappa\lambda}$, the $a_{\kappa\lambda\mu}$, and the characters, and which deserve some interest. Using (9) we derive from (64)

$$\varphi_\nu^{(\kappa)} \eta_\nu^{(\mu)} = \sum_\lambda c_{\mu\lambda} \varphi_\nu^{(\kappa)} \varphi_\nu^{(\lambda)} = \sum_{\lambda,\rho} c_{\mu\lambda} a_{\kappa\lambda\rho} \varphi_\nu^{(\rho)}. \tag{73}$$

We first set $\mu = \kappa$, and add over κ. By (16) we find

$$\delta_{\nu\nu^*} g/g_\nu = \sum_{\kappa,\lambda,\rho} c_{\kappa\lambda} a_{\kappa\lambda\rho} \varphi_\nu^{(\rho)}.$$

Here, we multiply by $g_\nu \eta_{\nu^*}^{(\sigma)}$, add over ν, and use (20)

$$\sum{}' \eta_\nu^{(\sigma)} = \sum_{\kappa,\lambda} c_{\kappa\lambda} a_{\kappa\lambda\sigma} \tag{74}$$

where the sum on the left extends over those ν for which the class $\mathfrak{C}_\nu$ is self-reciprocal, $\mathfrak{C}_\nu = \mathfrak{C}_{\nu^*}$.

Secondly, we take $\mu = \kappa'$ in (73) and apply the same method. We thus obtain

$$\sum_{\nu=1}^{k} \eta_\nu^{(\sigma)} = \sum_{\kappa,\lambda} c_{\kappa\lambda} a_{\kappa'\lambda\sigma}. \tag{75}$$

It can easily be seen from (67) that the number of self-contragredient modular characters, $\varphi^{(\lambda)} = \varphi^{(\lambda')}$ is equal to the number of self-reciprocal p-regular classes, $\mathfrak{C}_\nu = \mathfrak{C}_{\nu^*}$ $(\nu \leqq k)$.

24. Upper and lower bounds for the degrees of the indecomposable constituents U_κ. The product $\eta^{(\mu)} \cdot \varphi^{(\lambda')}$ can be expressed as a linear combination of the $\varphi^{(\rho)}$ (cf. (73)), and also, using (11), as a linear combination of the $\eta^{(\kappa)}$

$$\eta^{(\mu)} \cdot \varphi^{(\lambda')} = \sum \tilde{a}_{\kappa\lambda\mu} \eta^{(\kappa)}.$$

Here it is neither obvious that the coefficients are integers, nor that they are $\geqq 0$, but both these facts will follow from (73). Using (20) we find

$$g\tilde{a}_{\kappa\lambda\mu} = \sum_\nu g_\nu \eta_\nu^{(\mu)} \varphi_\nu^{(\lambda')} \varphi_\nu^{(\kappa')}$$

and comparing this with (70), and taking (68) into account, we obtain $\tilde{a}_{\kappa\lambda\mu} = a_{\kappa'\lambda'\mu'} = a_{\kappa\lambda\mu}$. Hence[30]

$$\eta^{(\mu)} \cdot \varphi^{(\lambda')} = \sum_\kappa a_{\kappa\lambda\mu} \eta^{(\kappa)}. \tag{76}$$

[30] This formula seems to indicate that $U_\mu \times F_{\lambda'}$ splits completely into $U_1, U_2, \cdots, U_k$ where U_κ appears $a_{\kappa\lambda\mu}$ times, but we have not been able to prove this.

In particular, we have $a_{\kappa\kappa'1} \geqq 1$, since $F_\kappa \times F_{\kappa'}$ contains the 1-representation. Hence $\eta^{(\kappa)}$ will appear in $\eta^{(1)} \cdot \varphi^{(\kappa)}$. On comparing the degrees, we find $u_1 f_\kappa \geqq u_\kappa$. On the other hand $a_{1\kappa\kappa} = 1$ and hence $\eta^{(1)}$ will appear in $\eta^{(\kappa)} \cdot \varphi^{(\kappa')}$. Consequently $u_1 \leqq f_\kappa u_\kappa$.

THEOREM 8. *For the degrees f_κ of the F_κ and u_κ of the U_κ there hold the inequalities*

$$u_1 f_\kappa \geqq u_\kappa \geqq u_1 / f_\kappa. \tag{77}$$

Since $u_\kappa = \sum_\lambda c_{\kappa\lambda} f_\lambda$ (cf. (12)), it follows from (77) that

$$u_\kappa \geqq c_{\kappa\kappa} f_\kappa + c_{\kappa\lambda} f_\lambda \qquad (\kappa \neq \lambda)$$

$$c_{\kappa\lambda} \leqq (u_1 - c_{\kappa\kappa}) \frac{f_\kappa}{f_\lambda} \quad \text{for} \quad \kappa \neq \lambda. \tag{78}$$

In particular, $c_{\kappa\lambda} < u_1$ since we may assume $f_\lambda \geqq f_\kappa$, further $c_{\kappa\kappa} \leqq u_1$.

On multiplying (77) by f_κ and adding, we find

$$u_1 \sum_{\kappa=1}^{k} f_\kappa^2 \geqq \sum_{\kappa=1}^{k} u_\kappa f_\kappa \geqq k u_1$$

The middle term here is g, as follows from (16). If the radical of the modular group ring Γ has the order m, then $\sum f_\kappa^2 = g - m$. Hence

$$u_1(g - m) \geqq g \geqq k u_1$$

$$g/k \geqq u_1 \geqq g/(g - m). \tag{79}$$

The multiplication of characters is used to obtain new characters if some characters have already been found. It is often convenient to determine the $\eta^{(\kappa)}$ at the same time with the $\varphi^{(\kappa)}$. Here formula (76) can be used. Formulas (77) and (79) can sometimes be used, if we want to show that a character η, which we have obtained, is an $\eta^{(\kappa)}$ and not a sum of several such $\eta^{(\kappa)}$.

V. RELATIONS BETWEEN THE CHARACTERS OF A GROUP $\mathfrak{G}$ AND THOSE OF A SUBGROUP $\mathfrak{H}$

25. The induced character. The second important method of Frobenius for the construction of characters assumes that the character χ of a representation V of a subgroup $\mathfrak{H}$ of $\mathfrak{G}$ is known. This representation "induces" a representation V^* of $\mathfrak{G}$ whose character χ^* can be obtained. The method of forming V^* remains valid in case we start with a modular representation, and so does the formula for χ^*, but this last formula requires a somewhat different proof here, due to the modified definition of the character of a representation.

Let h be the order of $\mathfrak{H}$, and let Q_μ ($\mu = 1, 2, \cdots, m$; $m = g/h$) be a complete residue system of $\mathfrak{G}$ (mod $\mathfrak{H}$)

$$\mathfrak{G} = \mathfrak{H}Q_1 + \mathfrak{H}Q_2 + \cdots + \mathfrak{H}Q_m.$$

We set $V(G) = 0$ if G does not belong to $\mathfrak{H}$ so that $F(G)$ is defined for all elements of $\mathfrak{G}$, and define

$$V^*(G) = (V(Q_\kappa G Q_\lambda^{-1})) \qquad (\kappa \text{ row index, } \lambda \text{ column index}). \tag{80}$$

It is easily seen that this is a representation V^* of $\mathfrak{G}$ of degree $tm = tg/h$ where t is the degree of V.

We shall determine the character of V^*. Let G be an arbitrary element of $\mathfrak{G}$. The element $Q_\mu G$ belongs to some residue class $\mathfrak{H}Q_{\rho_G(\mu)}$ and the permutations $P_G: \mu \to \rho_G(\mu)$ ($\mu = 1, 2, \cdots, m$) form a representation $P_{\mathfrak{G}}$ of $\mathfrak{G}$.

We split P_G into cycles. The length of each cycle is a divisor of the order of G. If G is p-regular, then the length of each cycle is prime to p. Let, for instance, $(1, 2, \cdots, \alpha)$ be the first cycle of P_G for a p-regular element G. Then $V^*(G)$ breaks up completely into the matrix

$$W = \begin{pmatrix} 0 & V(Q_1GQ_2^{-1}) & 0 & \cdots & 0 \\ 0 & 0 & V(Q_2GQ_3^{-1}) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & V(Q_{\alpha-1}GQ_\alpha^{-1}) \\ V(Q_\alpha GQ_1^{-1}) & 0 & 0 & \cdots & 0 \end{pmatrix}$$

and analagous matrices, corresponding to the other cycles of P_G. We have to determine the characteristic equation of W. We set $W_\nu = V(Q_\nu GQ_{\nu+1}^{-1})$ ($\nu = 1, 2, \cdots, \alpha - 1$), $W_\alpha = V(Q_\alpha GQ_1^{-1})$. We multiply the νth column of (81) on its right side by $W_{\nu-1}^{-1}W_{\nu-2}^{-1} \cdots W_1^{-1}$, and the νth row on its left side by $W_1W_2 \cdots W_{\nu-1}$ ($\nu = 2, 3, \cdots, \alpha$). Then $W_1, W_2, \cdots, W_{\alpha-1}$ in (81) are replaced by the unit matrix, whereas we have

$$\begin{aligned} W_1W_2 \cdots W_\alpha &= V(Q_1GQ_2^{-1})V(Q_2GQ_3^{-1}) \cdots V(Q_{\alpha-1}GQ_\alpha^{-1})V(Q_\alpha GQ_1^{-1}) \\ &= V(Q_1G^\alpha Q_1^{-1}) \end{aligned}$$

at the place of W_α in the last row, first column. Obviously, our changing of W amounts to a similarity transformation so that the characteristic polynomial remains unaltered. For the new form of W, the characteristic polynomial can be easily computed, and we find $f(x^\alpha)$ where $f(x)$ denotes the characteristic polynomial of $Q_1G^\alpha Q_1^{-1} = H$. As product of the elements $Q_\nu GQ_{\nu+1}^{-1}$ ($\nu = 1, 2, \cdots, \alpha - 1$) and $Q_\alpha GQ_1^{-1}$, this element H belongs to $\mathfrak{H}$. It is p-regular, since it is conjugate to G^α in $\mathfrak{G}$. Hence we obtain the characteristic roots of W by taking all the α^{th} roots of the characteristic roots of $V(H)$. This is still valid if we replace the characteristic roots (which lie in the modular field $\bar{K}$) by the corresponding complex roots of unity. Since α and the order of H are prime to p, no difficulty arises. It follows easily that if $\alpha > 1$ then the sum of these complex roots of unity must vanish. If $\alpha = 1$, then $W = V(Q_1GQ_1^{-1})$, and the sum in this case is the character $\chi(Q_1GQ_1^{-1})$. Dealing with the other cycles of P_G in the same manner, we obtain $\sum \chi(Q_\mu GQ_\mu^{-1})$ for the value of the character $\chi^*(G)$

of V^*. Here μ ranges over all values for which $Q_\mu G Q_\mu^{-1}$ lies in $\mathfrak{H}$. If we set $\chi(G) = 0$ for elements G outside of $\mathfrak{H}$ we may write

$$\chi^*(G) = \sum_{\mu=1}^{m} \chi(Q_\mu G Q_\mu^{-1}). \tag{82}$$

This is exactly the formula of Frobenius, only the trace argument by which it is ordinarily derived from (80) could not be used because of our modification in the definition of the modular characters (§6).

The splitting of the regular representation of $\mathfrak{H}$ into its indecomposable constituents corresponds to a decomposition of the group ring of $\mathfrak{H}$ into a direct sum of right ideals $\mathfrak{f}_\rho$. The products $h_\rho Q_\mu$ ($\mu = 1, 2, \cdots, m$) of Q_μ with the elements of $\mathfrak{f}_\rho$ generate a right ideal $\mathfrak{f}_\rho^*$ of the group ring Γ, and Γ is the direct sum of the $\mathfrak{f}_\rho^*$. If $\mathfrak{f}_\rho$ corresponds to the representation V_ρ of $\mathfrak{H}$ then it is known that $\mathfrak{f}_\rho^*$ corresponds to the representation V_ρ^* of $\mathfrak{G}$. It follows that the regular representation of $\mathfrak{G}$ breaks up completely into the V_ρ^*, corresponding to the different values of ρ. Each V_ρ^* itself consists of one or several of the indecomposable constituents U_κ.

We use for $\mathfrak{H}$ the same notations as for $\mathfrak{G}$ but with a $\sim$ sign, so $\tilde{F}_\kappa^*$ are the modular irreducible representations of $\mathfrak{H}$ etc. It follows now that $\tilde{U}_\kappa^*$ breaks up completely into some of the U_λ.

26. The formulas of Nakayama. We now can prove easily that we have formulas[31]

$$\begin{cases} \tilde{\eta}^{(\kappa)*} = \sum_\lambda \alpha_{\kappa\lambda}\eta^{(\lambda)} & \text{(for } p\text{-regular elements of } \mathfrak{G}) \\ \varphi^{(\lambda)} = \sum_\kappa \alpha_{\kappa\lambda}\tilde{\varphi}^{(\kappa)} & \text{(for } p\text{-regular elements of } \mathfrak{H}) \end{cases} \tag{83}$$

where the $\alpha_{\kappa\lambda}$ are rational integers, $\alpha_{\kappa\lambda} \geqq 0$. Obviously, the only point which requires a proof is that the same coefficients $\alpha_{\kappa\lambda}$ appear in both formulas. But from the orthogonality relations for the modular group characters (cf. (16), (20)) it follows for the coefficients $\alpha_{\kappa\lambda}$ of the first equation (83) that

$$g\alpha_{\kappa\lambda} = \sum \tilde{\eta}^{(\kappa)*}(G)\varphi^{(\lambda)}(G^{-1})$$

where the sum extends over all the p-regular elements G of $\mathfrak{G}$. Using (82) we obtain

$$g\alpha_{\kappa\lambda} = \sum_G \sum_{\mu=1}^{m} \tilde{\eta}^{(\kappa)}(Q_\mu G Q_\mu^{-1})\varphi^{(\lambda)}(G^{-1})$$

and after a simple rearrangement of the terms

$$h\alpha_{\kappa\lambda} = \sum_H \tilde{\eta}^{(\kappa)}(H)\varphi^{(\lambda)}(H^{-1})$$

[31] The formulas (83) and (85) are equivalent to those given in theorem 9 of T. Nakayama, Ann. of Math. 39, 1938, p. 361.

where H ranges over all the p-regular elements of $\mathfrak{H}$. This shows that $\alpha_{\kappa\lambda}$ is exactly the coefficient appearing in the second formula (83). This is exactly the method by which it is shown for ordinary characters that we have formulas

$$(84)\qquad \begin{cases} \tilde{\zeta}^{(i)*} = \sum_j l_{ij}\zeta^{(j)} & \text{(for elements of } \mathfrak{G}) \\ \zeta^{(j)} = \sum_i l_{ij}\tilde{\zeta}^{(i)} & \text{(for elements of } \mathfrak{H}) \end{cases}$$

where the l_{ij} are rational integers, $l_{ij} \geqq 0$.

Finally, we have formulas

$$(85)\qquad \begin{cases} \tilde{\varphi}^{(\kappa)*} = \sum_\lambda \beta_{\kappa\lambda}\varphi^{(\lambda)} & \text{(for } p\text{-regular elements of } \mathfrak{G}) \\ \eta^{(\lambda)} = \sum_\kappa \beta_{\kappa\lambda}\tilde{\eta}^{(\kappa)} & \text{(for } p\text{-regular elements of } \mathfrak{H}) \end{cases}$$

with rational integers $\beta_{\kappa\lambda} \geqq 0$, as can be shown in the same manner, or also be derived from (83).

We set $A = (\alpha_{\kappa\lambda})$, $B = (\beta_{\kappa\lambda})$, $L = (l_{ij})$. On comparing (84), (85) and (83) we obtain

$$\tilde{D}B = LD, \qquad DA' = L'\tilde{D}$$
$$AC = \tilde{C}B,$$

where $\tilde{D}$, $\tilde{C}$ have the same significance for $\mathfrak{H}$ as D, C have for $\mathfrak{G}$.

The second formula (83) shows that $\alpha_{11} = 1$, $\alpha_{\kappa 1} = 0$ for $\kappa \neq 1$ if the index 1 always refers to the 1-representation. Hence $\eta^{(1)}$ appears in the character $\tilde{\eta}^{(1)*}$ of degree $\tilde{u}_1 \cdot g/h$. Hence

$$(86)\qquad u_1 \leqq \tilde{u}_1 g/h.$$

In particular, if $\mathfrak{H}$ has an order prime to p, we have $\tilde{u}_1 = 1$, and obtain

THEOREM 9. *The degree of the indecomposable constituent of the regular representation of $\mathfrak{G}$ which corresponds to the* 1*-representation, is at most equal to the index of the maximal subgroup $\mathfrak{H}$ of an order prime to p.*

In particular, if $\mathfrak{G}$ has a subgroup $\mathfrak{H}$ of index p^a, then $u_1 \leqq p^a$, and since u_1 is divisible by p^a, we have $u_1 = p^a$. It follows that in this case U_1 is identical with the representation of $\mathfrak{G}$ by permutations which corresponds to the subgroup $\mathfrak{H}$ of index p^a (cf. the remark at the end of §25).

In the general case, it follows from (83) that every $\eta^{(\lambda)}$ appears in at least one $\tilde{\eta}^{(\kappa)*}$. Hence

$$(87)\qquad u_\lambda \leqq (g/h)\cdot \max(\tilde{u}_\kappa).$$

If $\mathfrak{H}$ again has an order prime to p, then $\tilde{u}_\kappa = \tilde{z}_\kappa = \tilde{f}_\kappa$, and we have $u_\lambda \leqq (g/h) \max(\tilde{f}_\kappa)$.

27. On the converse of theorem 1. We consider now the case that $\mathfrak{H}$ is the Sylow subgroup of order p^a of $\mathfrak{G}$. The character $\tilde{\varphi}^{(1)*}$ has here the value

$g' = g/p^a$ for the unit element and the value 0 for all the other p-regular elements. Hence

$$\sum_{\nu=1}^{k} g_\nu \bar{\varphi}_\nu^{(1)*} \eta_{\nu^*}^{(\lambda)} = g' u_\lambda = g \frac{u_\lambda}{p^a}.$$

Using the orthogonality relations (16), we obtain u_λ/p^a as multiplicity of $\varphi^{(\lambda)}$ in $\bar{\varphi}^{(1)*}$

$$\bar{\varphi}^{(1)*} = \sum_{\lambda=1}^{k} \frac{u_\lambda}{p^a} \varphi^{(\lambda)}.\text{[32]} \tag{88}$$

From (84) we have

$$\bar{\zeta}^{(1)*} = \sum_i l_i \zeta^{(i)} \qquad (l_i = l_{1i}). \tag{89}$$

For p-regular elements, these two characters are identical. By (12) and (10) we have $u_\lambda = \sum d_{i\lambda} z_i$, $\zeta^{(i)} = \sum d_{i\lambda} \varphi^{(\lambda)}$, and thus obtain

$$\sum_\lambda \sum_i \frac{d_{i\lambda} z_i}{p^a} \varphi^{(\lambda)} = \sum \sum l_i d_{i\lambda} \varphi^{(\lambda)}.$$

for all p-regular elements of $\mathfrak{G}$. Hence, on comparing the coefficients of $\varphi^{(\lambda)}$

$$\sum_i \left(\frac{z_i}{p^a} - l_i\right) d_{i\lambda} = 0 \qquad (\lambda = 1, 2, \cdots, k). \tag{90}$$

Assume now that we have a block B_τ which contains the same number of ordinary and modular representations, $x_\tau = y_\tau$. We choose the index λ in (90) such that $\varphi^{(\lambda)}$ belongs to B_τ. Then it is sufficient to let i range over those values for which $\zeta^{(i)}$ belongs to B_τ, since for other values of i, $d_{i\lambda} = 0$. We may consider (90) as a system of y_τ linear homogeneous equations for the $x_\tau = y_\tau$ quantities $z_i/p^a - l_i$. Since the determinant is D'_τ and has the rank y_τ, all the $(z_i/p^a) - l_i$ vanish. But l_i is an integer, hence $z_i \equiv 0 \pmod{p^a}$. This shows again, that if $x_\tau = y_\tau$ then B_τ is a block of highest kind (converse of the middle part of theorem 1, cf. §14).

Using (84), we see that l_i in (89) is also the coefficient with which $\bar{\zeta}^{(1)}$ appears in $\zeta^{(i)}$. Hence

$$l_i = \frac{1}{p^a} \sum \zeta^{(i)}(H) \tag{91}$$

where H ranges over all the elements of the Sylow group $\mathfrak{H}$ of order p^a. The first term here is z_i/p^a. Combining this formula with the results of §22, we easily can obtain (90) again.

The formula (91) shows that if $\zeta^{(i)}(H)$ vanishes for all elements of an order

[32] This formula shows Dickson's theorem, $p^a \mid u_\lambda$ (cf. footnote 6), and this is essentially the way by which Dickson proved his result.

p^μ with $\mu > 0$, then $l_i = z_i/p^a$. Hence in this case $\zeta^{(i)}$ must be a character of the highest kind.

THEOREM 10. *If an irreducible character $\zeta^{(i)}$ vanishes for all elements of an order p^μ, $\mu > 0$, then $\zeta^{(i)}$ is a character of the highest kind.*

This is a converse to the last part of theorem 1. It even would be sufficient to assume that all the $\zeta^{(i)}(H)$ are divisible by p^a, if H is an element of order p^μ, $\mu > 0$. If all these $\zeta^{(i)}(H)$ are divisible by p^a, then $p^a \mid z_i$. This result can easily be improved when we take into account the multiplicity with which the terms $\zeta^{(i)}(H)$ appear on the right side of (91).

28. An upper and a lower bound for c_{11}. We now consider two arbitrary subgroups $\mathfrak{H}$ and $\mathfrak{J}$ of $\mathfrak{G}$ of orders h and j. Let $\alpha(G)$ and $\beta(G)$ be the ordinary characters of $\mathfrak{G}$, induced by the 1-representations of $\mathfrak{H}$ and $\mathfrak{J}$ respectively. From (82) we obtain easily that $\alpha(G)h$ is the number of elements M in $\mathfrak{G}$ for which G lies in $M^{-1}\mathfrak{H}M$, and similarly $\beta(G)j$ is the number of elements N for which G lies in $N^{-1}\mathfrak{J}N$. Then G will lie in $hj\alpha(G)\beta(G)$ of the intersections $\mathfrak{D}_{M,N} = [M^{-1}\mathfrak{H}M, N^{-1}\mathfrak{J}N]$.

If $\mathfrak{D}_{M,N}$ has the order $t_{M,N}$ then

$$(92) \qquad hj \sum_G \alpha(G)\beta(G) = \sum_{M,N} t_{M,N} = \sum_{M,N} t_{MN^{-1},1} = g \sum t_{M,1}.$$

On the other hand, we split $\mathfrak{G}$ into residue classes modd $\mathfrak{H}$ and $\mathfrak{J}$.

$$(93) \qquad \mathfrak{G} = \sum_{\nu=1}^{r} \mathfrak{H}R_\nu\mathfrak{J}.$$

The number of elements in $\mathfrak{H}R_\nu\mathfrak{J}$ is equal to $hj/t_{M,1}$ where M is any element of $\mathfrak{H}R_\nu\mathfrak{J}$. Hence, if M ranges over the elements of $\mathfrak{H}R_\nu\mathfrak{J}$, $\sum' t_{M,1} = hj$, and from (92) it follows that

$$(94) \qquad \begin{aligned} hj \sum_G \alpha(G)\beta(G) &= g\cdot rhj \\ \sum_G \alpha(G)\beta(G) &= g\cdot r \end{aligned}$$

where r is the number of residue classes of $\mathfrak{G}$ (modd $\mathfrak{H}$, $\mathfrak{J}$).

We assume now that $\mathfrak{H}$ and $\mathfrak{J}$ have orders prime to p. We may restrict the summation on the left hand of (94) to p-regular elements since for the other elements $\alpha(G) = 0$, and may consider $\alpha(G)$ and $\beta(G)$ as the modular characters of $\mathfrak{G}$, induced by the 1-representations of $\mathfrak{H}$ and $\mathfrak{J}$. We may set

$$\alpha(G) = \sum a_\kappa \eta^{(\kappa)}(G), \qquad \beta(G) = \sum b_\lambda \eta^{(\lambda)}(G)$$

where according to (83), a_κ and b_λ are rational integers ≥ 0, and $a_1 = 1$, $b_1 = 1$.

$$\alpha(G) = \sum_{\kappa,\lambda} a_\kappa c_{\kappa\lambda} \varphi^{(\lambda)}(G).$$

It is easily verified by use of (82) that $\beta(G) = \beta(G^{-1})$ and so combining (94) and the orthogonality relations (20), we obtain

$$\sum_{\kappa,\lambda=1}^{k} a_\kappa c_{\kappa\lambda} b_\lambda = r. \tag{95}$$

In particular,

$$c_{11} \leqq r. \tag{96}$$

THEOREM 11. *If $\mathfrak{H}$ and $\mathfrak{J}$ are two subgroups of $\mathfrak{G}$, whose orders are prime to p, then the first Cartan invariant c_{11} is at most equal to the number of residue classes of $\mathfrak{G}$* (mod $\mathfrak{H}$, $\mathfrak{J}$).

If, for instance, $\mathfrak{G}$ is a doubly transitive permutation group of order p^a, then we may take $\mathfrak{H} = \mathfrak{J}$ as subgroup of index p^a. Here $r = 2$, and hence $c_{11} = 2$.

Since C is the matrix of a positive definite quadratic form, the coefficient in the first row and first column of C^{-1} is at least $\frac{1}{c_{11}}$, and the equality sign is possible only if $c_{1s} = 0$ for all $s > 1$, i.e. if the modular 1-representation forms a block of its own. In this exceptional case, we have $C_1 = (c_{11})$, $c_{11} = p^a$, since p^a is the only elementary divisor of C_1. Then $\mathfrak{G}$ contains a normal subgroup $\mathfrak{H}$ of index p^a, (cf. §29 below).

In any case, we find from $gC^{-1} = (\gamma_{\kappa\lambda})$ and (35)

$$N = \tilde{\gamma}_{11} \geqq g/c_{11}$$

$$c_{11} \geqq g/N. \tag{97}$$

THEOREM 12. *The first Cartan invariant c_{11} is at least equal to g/N where N is the number of elements of an order prime to p in $\mathfrak{G}$. The equality sign holds only if these elements form a normal subgroup, necessarily of index p^a.*

From (96) and (97) it follows that $rN > g$, except for the case that $\mathfrak{G}$ contains a normal subgroup of index p^a. If, for instance, g is divisible by three distinct primes $g = p^a p_1^{a'} p_2^{a''}$, then we can take for $\mathfrak{H}$ and $\mathfrak{J}$ the Sylow-groups of orders $p_1^{a'}$, and $p_2^{a''}$ and have $r = p^a$. Hence $N > g/p^a$, except when $\mathfrak{G}$ contains a normal subgroup of index p^a.[33]

VI. Special Cases and Examples

29. Special cases. We first consider the case that $\mathfrak{G}$ is a direct product, $\mathfrak{G} = \mathfrak{A} \times \mathfrak{B}$. If $A \to F(A)$ is a representation of $\mathfrak{A}$, and $B \to K(B)$ is a representation of $\mathfrak{B}$, then $A \times B \to F(A) \times K(B)$ (Kronecker product) is a representation of $\mathfrak{A} \times \mathfrak{B}$. This representation $F \times K$ is irreducible, if F and K are irreducible, and conversely, every irreducible representation of $\mathfrak{A} \times \mathfrak{B}$ is of

[33] This is a very special case of an unproved conjecture of Frobenius which states that when there are exactly r elements X an order dividing r in a group of these elements form a subgroup.

this form.[34] This implies that the D-matrix of $\mathfrak{G}$ is the direct product of the D-matrices of $\mathfrak{A}$ and of $\mathfrak{B}$, and that the C-matrix of $\mathfrak{G}$ is the product of the C-matrices of $\mathfrak{A}$ and of $\mathfrak{B}$.

We next consider the case that $\mathfrak{G}$ contains a normal subgroup $\mathfrak{H}$ of index $p^a (g = p^a g'$ with $(g', p) = 1)$. Since $\eta^{(1)} = \bar{\varphi}^{(1)*}$ (cf. §26). we see that $\eta^{(1)}$ has the value p^a for every p-regular element. It follows that $\eta^{(1)} = p^a\varphi^{(1)}$, $C_1 = (p^a)$, and $\varphi^{(1)}$ is the only irreducible modular character in the first block B_1. Conversely, let $\mathfrak{G}$ be a group for which the first block B_1 contains only one irreducible modular character. Denote by $\mathfrak{H}$ the normal subgroup whose elements are represented by the unit matrix I in each ordinary irreducible representation Z_i of the first block. Then Z_i is a representation of $\mathfrak{G}/\mathfrak{H}$, and the index g/h is at least equal to the sum of the squares of the degrees z_i of these Z_i. This sum is equal to $u_1 f_1 \geqq p^a$, hence $g/h \geqq p^a$. On the other hand, each p-regular element of $\mathfrak{G}$ is represented in Z_i by a matrix whose characteristic roots are all 1, and which then is equal to I. This shows that h is divisible by every prime power dividing g/p^a. Hence $h \geqq g/p^a$, so $h = g/p^a$. We see that $\mathfrak{G}$ contains a normal subgroup of index p^a, if and only if, the first block contains only one irreducible modular character.

Let us assume now that $\mathfrak{G}$ contains a normal subgroup $\mathfrak{P}$ of order p^a, $(g = p^a g', (g', p) = 1)$. Every representation of $\mathfrak{G}/\mathfrak{P}$ defines a representation of $\mathfrak{G}$. In particular, the regular representation of $\mathfrak{G}/\mathfrak{P}$ has the character χ, $\chi(1) = g/p^a$, $\chi(G) = 0$ for a p-regular element $G \neq 1$. This is the character $\bar{\varphi}^{(1)*}$ of §27. From (88), it follows that this character contains $\varphi^{(\lambda)}$ exactly u_λ/p^a times. In particular, $\varphi^{(\lambda)}$ must represent the elements of $\mathfrak{P}$ by the unit matrix.[35] Further, since the regular representation of $\mathfrak{G}/\mathfrak{P}$ contains each of its irreducible constituents $\varphi^{(\lambda)}$ exactly f_λ times, we have $f_\lambda = u_\lambda/p^a$. The degree f_λ is prime to p, since $\varphi^{(\lambda)}$ is a representation of $\mathfrak{G}/\mathfrak{P}$ of order g', so each block is of the lowest kind. We do not know whether the converse is true.

Finally, let $\mathfrak{G}$ be a group in which every p-regular element commutes with every element of a p-Sylow group. Then we have k blocks of the lowest kind. Each of them can contain only one modular irreducible constituent, and C necessarily has the form

$$C = \begin{pmatrix} p^a & & & \\ & p^a & 0 & \\ & & \ddots & \\ & 0 & & p^a \end{pmatrix}.$$

The converse is also true.

In the same manner as for ordinary characters, it follows that every linear character of any group $\mathfrak{G}$ is actually a character of $\mathfrak{G}/\mathfrak{K}$ where $\mathfrak{K}$ denotes the

[34] The first part follows easily from Burnside's theorem, Proc. London Math. Soc. (2) 3, 1905, p. 430; the second from A. H. Clifford's theorem, Ann. of Math. 38, 1937, p. 533.

[35] This follows from the fact that F_λ appears as a constituent in a representation for which this is true.

commutator group of $\mathfrak{G}$. Hence, the number of linear modular characters of $\mathfrak{G}$ is equal to the largest factor of the index of the commutator group which is prime to p. For a linear character $\varphi^{(\lambda)}$, the formula (77) shows that $u_\lambda = u_1$. This can also be shown directly, as U_λ may be expressed as the direct product of U_1 and $\varphi^{(\lambda)}$ (cf. (76), (64)). The linear character $\varphi^{(\lambda)}$ will belong to the block B_1, if and only if $\varphi^{(\lambda)} = 1$ for every p-regular element G which commutes with every element of a p-Sylow group. The block B_τ of $\varphi^{(\lambda)}$ is obtained from the block B_1 by multiplication with $\varphi^{(\lambda)}$.

30. The groups $GLH(2, p^a)$, $SLH(2, p^a)$, **and** $LF(2, p^a)$. As first examples we treat the group $SLH(2, p^a)$ of all matrices $\begin{pmatrix} m_{11}, & m_{12} \\ m_{21}, & m_{22} \end{pmatrix} = G$ of determinant 1 with coefficients in the Galois field $GF(p^a)$ with $p^a = q$ elements. The rth power matrices $G^{(r)}$ form a representation $\mathfrak{G}^{(r)}$ for any fixed $r = 0, 1, 2, \cdots$. Further, if $\mathfrak{H}$ is a modular representation with coefficients in $GF(q)$, we obtain a new representation by applying an automorphism θ of $GF(q)$ to all coefficients; we denote this representation by $\mathfrak{H}^\theta$. Let θ_ν now be the automorphism $\alpha \to \alpha^{p^\nu}$ of $GF(q)$ $(\nu = 1, 2, \cdots, a-1)$. We form the representation

$$H(r_0, r_1, \cdots, r_{a-1}) = \mathfrak{G}^{(r_0)} \times \mathfrak{G}^{(r_1)\theta_1} \times \cdots \times \mathfrak{G}^{(r_{a-1})\theta_{a-1}} \tag{98}$$

$(r_\nu = 0, 1, 2, \cdots, p-1)$. We thus obtain p^a modular representations, and state that these are all the irreducible representations. In order to prove this, we first notice that there are exactly p^a classes of p-regular conjugate elements in $\mathfrak{G}$ since each such class corresponds in a $(1-1)$ manner to a polynomial $x^2 - tr(G)\,x + 1$, the characteristic polynomial of its elements, and $tr(G)$ can be any element of $GF(q)$. Secondly, we prove that the representations (98) are irreducible. Let x_ν, y_ν undergo the transformation G^{θ_ν}, $(\nu = 0, 1, 2, \cdots, a-1)$. Then (98) belongs to the vector-module $\mathfrak{B}$ of all polynomials in x_0, $y_0, \cdots, x_{a-1}, y_{a-1}$ which are homogeneous of degree r_ν in x_ν, y_ν, $(\nu = 0, 1, \cdots, a-1)$.[36] We have to show that $\mathfrak{B}$ is irreducible, when the elements of $\mathfrak{G}$ are taken as operators. If F now is any element of $\mathfrak{B}$, the module $\mathfrak{M}(F)$ generated by F will contain all the polynomials F_t which are obtained from F by applying $x \to x + ty$, $y \to y$ for any t in $GF(q)$. Then $x_\nu \to x_\nu + t^{p^{\nu-1}} y_\nu$, $y_\nu \to y_\nu$. Obviously, F_t is of the form

$$F_t = H_0 + tH_1 + \cdots + t^{p^a-1} H_{p^a-1},$$

where H_μ depends on x_ν, y_ν. We now take for t the q different elements of $GF(q)$. It follows easily that each H_μ is a linear combination of the F_t, and hence lies in $\mathfrak{M}(F)$. The last H_μ which is not zero obviously is a single power product $Ay_0^{r_0}y_1^{r_1} \cdots y_{a-1}^{r_{a-1}}$, and hence $y_0^{r_0}y_1^{r_1} \cdots y_{a-1}^{r_{a-1}}$ lies in $\mathfrak{M}(F)$. We replace F now by this polynomial, apply the transformation $x \to x$, $y \to tx + y$, and use the same argument. We thus see that every power product of $x_0, y_0, \cdots, x_{a-1}$, y_{a-1} of the correct degrees lies in $\mathfrak{M}(F)$. Hence $\mathfrak{M}(F) = \mathfrak{B}$, i.e. (98) is irreducible.

[36] The coefficients of these polynomials can be taken from any extension field of $GF(q)$.

Finally, the representations (98) are all distinct. In order to show that, assume

$$H(r_0, r_1, \cdots, r_{a-1}) = H(r_0', r_1', \cdots, r_{a-1}') \tag{99}$$

$0 \leqq r_i \leqq p-1$, $0 \leqq r_i' \leqq p-1$, and $r_i = r_i'$ does not hold for all i. We arrange the $H(r_0, r_1, \cdots, r_{a-1})$ in lexicographical order by taking $H(r_0, r_1, \cdots, r_{a-1})$ as lower than $H(s_0, s_1, \cdots, s_{a-1})$ when the first difference $s_0 - r_0$, $s_1 - r_1$, $\cdots$ $s_{a-1} - r_{a-1}$ which does not vanish, has a positive value. We may assume that $H(r_0, r_1, \cdots, r_{a-1})$ is the lowest representation (98) which is similar to another of these representations. Certainly not all the r can be equal to $p - 1$. But if all the r_ν' were equal to $p - 1$ then the right side in (99) would have the maximum degree p^a which would imply that all the $r_i = p - 1 = r_i'$. This case is, therefore, also excluded.

Assume $r_0 = \cdots = r_{i-1} = 0$, $r_i \neq 0$, $i \geqq 0$. We multiply (99) by $\mathfrak{G}^{\theta_\nu}$ (Kronecker product). We can express both sides as sums of representations (98) again when we use repeatedly the relations

$$\mathfrak{G}^0 \times \mathfrak{G}^{\theta_\nu} = \mathfrak{G}^{\theta_\nu}, \quad \mathfrak{G}^{(r)\theta_\nu} \times \mathfrak{G}^{\theta_\nu} \leftrightarrow \mathfrak{G}^{(r-1)\theta_\nu} + \mathfrak{G}^{(r+1)\theta_\nu}, \qquad (r = 1, 2, \cdots, p-1)$$

$$\mathfrak{G}^{(p-1)\theta_\nu} \times \mathfrak{G}^{\theta_\nu} \leftrightarrow \mathfrak{G}^{\theta_{\nu+1}} + 2\mathfrak{G}^{(p-2)\theta_\nu}.$$

After the multiplication, $H(r_0, \cdots, r_{i-1}, r_i - 1, r_{i+1}, \cdots, r_{a-1})$ will appear on the left side of (99), whereas this term cannot appear on the right side. Because of the uniqueness of the irreducible constituents, we obtain a contradiction. This shows that the representations (98) are all the modular irreducible representations of $\mathfrak{G}$.

In the case of $GLH(2, p^a)$, we have to add a factor Δ^s, $0 \leqq s \leqq q - 2$, $\Delta = m_{11}m_{22} - m_{12}^2$, on the right side of (98), in order to obtain all the irreducible modular representations.

On the other hand, if (98) is to give a representation of the factor group $LF(2, p^a)$ of $SLH(2, p^a)$ modulo its centrum, then $-I$ must be represented by I in (98), i.e. the number $r_0 + r_1 + r_2 + \cdots + r_{a-1}$ must be even.[37]

31. The Cartan invariants and decomposition numbers (mod p) **of** $LF(2, p)$. We restrict ourselves to the case $LF(2, p)$, p an odd prime. The irreducible modular characters are here $\mathfrak{G}^{(0)}, \mathfrak{G}^{(2)}, \cdots, \mathfrak{G}^{(p-1)}$, the degrees are $1, 3, \cdots, p$. This shows, in particular, that the degree of the irreducible modular representations need not be a divisor of the order of the group. For the order of the radical we obtain

$$g - 1^2 - 3^2 - \cdots - p^2 = \frac{p(p^2-1)}{2} - \frac{(p+2)(p+1)p}{6} = \frac{p(p+1)(2p-5)}{6}.$$

[37] The modular characters of $GLH(3, p)$, $SLH(3, p)$ and $LF(3, p)$ have been determined by C. Mark in his Toronto thesis (to appear in the University of Toronto Studies).

There is no difficulty in computing the modular characters, and if they are arranged in the order $\mathfrak{G}^{(0)}, \mathfrak{G}^{(p-3)}, \mathfrak{G}^{(2)}, \mathfrak{G}^{(p-5)}, \cdots, \mathfrak{G}^{(p-3)/2}$ or $\mathfrak{G}^{(p-1)/2}, \mathfrak{G}^{(p-1)}$ we have

$$(100) \qquad C_1 = \begin{pmatrix} 2 & 1 & & & & \\ 1 & 2 & 1 & & & \\ & 1 & 2 & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot\ 2 & 1 \\ & & & & 1 & 3 \end{pmatrix} \qquad C_2 = (1)\ ^{38}$$

for the C parts corresponding to the two blocks. (The coefficients not filled in are 0, there is just one 3 in the main diagonal of C_1). The first and the last u_κ both have the value p, all the other u_κ have the value $2p$.

Using formula (100) we can find D without any ambiguity. There must be two 1's in the first column. Beside them, we must have a 0 and a 1 in the second column etc. We thus find

$$D_1 = \begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1\cdot & & & \\ & & \cdot\ \cdot & & \\ & & & \cdot\ \cdot & \\ & & & 1 & 1 \\ & & & & 1 \\ & & & & 1 \end{pmatrix} \qquad D_2 = (1)$$

In this manner, we may obtain the values of the ordinary characters of $LF(2, p)$ except for the two p-singular classes. There is no difficulty in obtaining these missing values.[39]

University of Toronto,
University of Michigan.

---

[38] The determination of C according to this method is rather complicated, and we do not give the details of the computation. Using the methods sketched in op. cit. in footnote 20, C and D can be determined easily.

[39] These characters were first given by G. Frobenius, Sitzungsber. Preuss. Akad. 1896, p. 1013; the ordinary characters of the binary groups in $GF(p^a)$, $a > 1$ were first given by I. Schur, Jour. reine angew. Math. 132, 85, 1907, and independently by H. E. Jordan, Am. Jour. of Math. 29, 1907, p. 387.